

**Issues related to the economic analysis of genomic
diagnostic technologies in the UK National Health
Service: an exploration of methods**

James Buchanan

**Oriel College
University of Oxford
Trinity Term 2015**

D.Phil. Thesis

Acknowledgements

This thesis represents the culmination of four years of work and research and would not have been possible without the guidance and support of many people. First and foremost I would like to thank my supervisors, Dr Sarah Wordsworth and Dr Anna Schuh, for providing invaluable feedback and constant encouragement, for challenging me to think and work outside my comfort zone, and for progressing my development as an academic researcher. My thanks also go to the UK National Institute for Health Research who provided funding for me to work on this thesis for most of the last four years.

I would also like to thank a number of health economists who provided advice regarding the discrete choice experiment (DCE), including Professor Katherine Payne, Emily Holmes, Caroline Vass, Dr Mark Harrison, Brendan Mulhern, Dr Verity Watson and Dr Arne Risa Hole. In addition, several health economists posed challenging questions related to my work at three Health Economists' Study Group meetings in Oxford, Sheffield and Glasgow, in particular Dr Emma Frew, and all of these comments improved this thesis.

This thesis addresses several research questions in health economics using a case study in chronic lymphocytic leukaemia (CLL). My thanks go to those who gave their time to help me to understand this clinical area, including Dr Samantha Knight, Ruth Clifford, Dr Susanne Weller, Dr Karthik Ramasamy, Dr Tim Littlewood, Dr George Follows and all the members of the Oxford Biomedical Research Centre Haemato-Molecular Diagnostic Laboratory who assisted with data collection. Special thanks go to the UK CLL Support Association, especially Chonette Taylor, for their assistance with the DCE, and to all the patients who completed a DCE survey.

Moral support is just as important as scientific support when undertaking a DPhil. Particular thanks go to Filipa, Anees, Liz and Reem who all embarked on this journey with me at the same time. The work in this thesis was completed in three different offices, in the company of numerous supportive office-mates: my thanks go to José, Ramón, Helen, Liz, Mara, Iryna, Claire and Rachael. In addition, Alastair, Ramón, Sam, Ingrid, José, Jane, Oliver and Helen all provided invaluable feedback on various chapters, which was very much appreciated.

I would like to thank my family for their lifelong love and support, without which I would not have even started, let alone completed, this journey. In addition, over the past four years I met, moved in with, got engaged to, and married my amazing wife, Sarah. Sarah has provided relentless support, encouragement and patience over this period, which has included a delayed honeymoon and multiple house-moves. Words cannot express how much I have appreciated having Sarah in my life over the past four years.

Finally, this thesis is dedicated to the memory of my Grandad Bob, who spent countless hours passing on his love of mathematics to me during his lifetime. He was fascinated by patterns in numbers, and Paul Erdos summed up his thoughts on the matter when, in response to the question ‘Why are numbers beautiful?’, he said:

“It’s like asking why is Ludwig van Beethoven’s Ninth Symphony beautiful. If you don’t see why, someone can’t tell you. I know numbers are beautiful. If they aren’t beautiful, nothing is.”

My Grandad planned on attending Queens University in Belfast but was unable to do so because of the onset of World War Two. By virtue of this mention in my thesis, he’ll now always be a part of the Bodleian Library. I think this would have made him very happy.

Publications and presentations related to this thesis

Publications

Buchanan J, Wordsworth S. Welfarism Versus Extra-Welfarism: Can the Choice of Economic Evaluation Approach Impact on the Adoption Decisions Recommended by Economic Evaluation Studies? *PharmacoEconomics* 2015;33(6):571-9.

Buchanan J, Wordsworth S, Schuh A. Issues surrounding the health economic evaluation of genomic technologies. *Pharmacogenomics* 2013;14(15):1833-47.

Presentations

Buchanan J, Wordsworth, S, Schuh A. [*Presentation JB*] Presenting risk information in discrete choice experiments: do graphical techniques influence willingness-to-pay? July 2015. International Health Economics Association 11th World Congress, Milan, Italy.

Buchanan J, Fermont J. [*Presentation JB/JF*] Methodological issues surrounding the health economic evaluation of genomic technologies and a case study of these issues in the research setting. March 2015. Personalised Medicine and Resource Allocation Conference, University of Oxford, UK.

Buchanan J, Wordsworth, S, Schuh A. [*Discussion paper*] Presenting risk information in discrete choice experiments: do graphical techniques influence willingness-to-pay? June 2014. Health Economists' Study Group Meeting, Glasgow Caledonian University, UK.

Buchanan J. Evaluating genomic technologies: Can health economics improve its methodological toolbox? January 2014. Academic Unit of Health Economics, University of Leeds, UK.

Buchanan J, Wordsworth S. [*Discussion paper*] Welfarism versus extra-welfarism: does choice of economic evaluation approach impact on adoption decisions? January 2014. Health Economists' Study Group Meeting, University of Sheffield, UK.

Buchanan J. Health economics, genomic testing and CLL. October 2013. CLL Support Association Members' meeting, St Edmund Hall, The University of Oxford, UK.

Buchanan J, Wordsworth S, Schuh A. [*Presentation JB*] Evaluating genomic technologies: Can health economics improve its methodological toolbox? July 2013. International Health Economics Association Conference 10th World Congress, Sydney, Australia.

Buchanan J, Wordsworth S, Taylor J, Schuh A, Knight S. [*Discussion paper*] Assessing genomic technologies: do we need to overhaul the health economists' methodological toolbox? June 2012. Health Economists' Study Group Meeting, University of Oxford, UK.

Buchanan J. Health economics, genetic testing and CLL. March 2012. CLL Support Association Members' meeting, The Oxford Belfry Hotel, Milton Common, UK.

Abstract

Issues related to the economic analysis of genomic diagnostic technologies in the UK National Health Service: an exploration of methods

James Buchanan, Oriel College.
D.Phil.
Trinity Term 2015.

This thesis explores the empirical and methodological issues that arise when conducting economic analyses of genomic technologies in the UK National Health Service. The main focus of the thesis is the identification of the most appropriate economic evaluation approach to use in this context. Health economists are currently uncertain as to whether existing extra-welfarist approaches are sufficient to evaluate genomic technologies. Clarification is required because alternative methods, such as the welfarist approach, may lead to different adoption decisions, which will impact on population health.

The advantages and disadvantages of alternative approaches are demonstrated using a clinical case study: the use of genomic testing to guide treatment decisions in chronic lymphocytic leukaemia (CLL). Two forms of extra-welfarist economic evaluation are conducted – a cost-effectiveness analysis (CEA) and a cost-utility analysis (CUA) – along with a welfarist economic evaluation which took the form of a cost-benefit analysis (CBA), informed by a discrete choice experiment (DCE). These analyses consider several test-treatment strategies, reflecting different combinations of testing and the use of both existing therapies and new treatments such as ibrutinib. Information is also generated on patient preferences in this context, via the DCE, and the results of a survey are presented in which stakeholders are asked to consider which form of economic evaluation best captures the true costs and benefits of genomic testing in CLL.

The results of the CEA, CUA and CBA all indicate that the optimal strategy is to not adopt genomic testing in this context. The different approaches are then compared to assess their relative merits. The key finding – which is supported by the stakeholder survey – is that the extra-welfarist approach to economic evaluation is most appropriate when assessing genomic interventions. However, an evolution in methods is required to ensure that the results of these analyses provide decision-makers with sufficient information to facilitate the efficient allocation of scarce healthcare resources in this context.

Table of contents

Acknowledgements	ii
Publications and presentations related to this thesis	iv
Abstract	vi
Table of contents.....	vii
List of tables.....	xiii
List of figures.....	xv
List of abbreviations	xvi
Chapter One Introduction	1
1.1 Rationale for the thesis	1
1.2 Clinical case study	5
1.3 Thesis aims and objectives	6
1.4 Structure of the thesis.....	7
1.5 Contribution to the literature	10
1.6 Conclusion	11
Chapter Two Conducting economic evaluations of genomic technologies: methodological challenges	12
2.1 Introduction.....	12
2.2 Background to economics and health economic evaluation	13
2.2.1 Rationale for economic evaluation	13
2.2.2 Types of economic evaluation.....	15
2.3 Welfarism versus extra-welfarism: which approach should be used?	18
2.3.1 Evolution of the debate over time	18
2.3.2 Can the choice of economic evaluation approach impact on the adoption decisions recommended by economic evaluation studies?	22
2.4 Methodological challenges in genomics	32
2.4.1 Literature review aims, objectives and methods	32
2.4.2 Literature review results	34
2.4.3 Discussion.....	52
2.5 Conclusions.....	55
Chapter Three Chronic lymphocytic leukaemia and genetic testing	57

3.1 Introduction.....	57
3.2 Chronic lymphocytic leukaemia	57
3.2.1 Epidemiology.....	57
3.2.2 Presentation, diagnosis and staging.....	58
3.2.3 Genetic testing in chronic lymphocytic leukaemia	59
3.2.4 Clinical management and prognosis	71
3.3 Targeted next generation sequencing	78
3.3.1 Overview	78
3.3.2 Targeted next generation sequencing in chronic lymphocytic leukaemia	80
3.3.3 Health economic evidence for the use of targeted next generation sequencing technology in all disease areas.....	81
3.3.4 Barriers to more widespread use of targeted next generation sequencing technology	83
3.4 Discussion and conclusions	84
Chapter Four An extra-welfarist economic evaluation of genomic testing in chronic lymphocytic leukaemia.....	86
4.1 Introduction.....	86
4.2 Selecting an analytical approach.....	87
4.3 Devising the model structure	90
4.3.1 Comparators considered in the model	92
4.3.2 Additional design considerations	98
4.4 Populating the model with transition probabilities, costs and outcomes	99
4.4.1 Genetic parameters.....	100
4.4.2 Transition probabilities	102
4.4.3 Resource use and unit costs.....	105
4.4.4 Health outcomes	111
4.5 Analytical methods.....	113
4.5.1 Base case analysis	113
4.5.2 Sensitivity and scenario analysis	114
4.6 Results	115
4.6.1 Base case analysis	116
4.6.2 Sensitivity and scenario analysis	118
4.6.3 Probabilistic sensitivity analysis	123
4.7 Discussion and conclusions	127
Chapter Five A welfarist economic evaluation of genomic testing in chronic lymphocytic leukaemia.....	131
5.1 Introduction.....	131

5.2 Selecting an analytical approach.....	132
5.2.1 Potential analytical approaches	132
5.2.2 Existing cost-benefit analyses which use willingness-to-pay estimates generated by a discrete choice experiment.....	137
5.2.3 Issues surrounding the presentation of risk information in discrete choice experiments	139
5.3 Designing the discrete choice experiment.....	142
5.3.1 Objectives.....	142
5.3.2 The design process	142
5.4 Discrete choice experiment results.....	152
5.4.1 Survey metrics and respondent characteristics	152
5.4.2 Model A results	154
5.4.3 Model B results	158
5.4.4 Impact of risk graphics on willingness-to-pay for genomic testing.....	160
5.4.5 Discussion.....	161
5.5 Designing the cost-benefit analysis.....	168
5.5.1 Monetising the benefits of genetic and genomic testing	169
5.5.2 Cost data used in the cost-benefit analysis.....	171
5.5.3 Data analysis.....	172
5.6 Cost-benefit analysis results	173
5.6.1 Main results.....	173
5.6.2 Sensitivity analysis and probabilistic sensitivity analysis	174
5.6.3 Discussion.....	176
5.7 Conclusions.....	178
Chapter Six Comparing the welfarist and extra-welfarist approaches to economic evaluation in chronic lymphocytic leukaemia.....	180
6.1 Introduction.....	180
6.2 Facilitating a quantitative comparison between the two approaches	181
6.2.1 Extending the extra-welfarist economic evaluation to a societal perspective	182
6.2.2 Extending the extra-welfarist economic evaluation to consider different time horizons	186
6.2.3 Undertaking value of information analysis	189
6.2.4 Adapting the welfarist economic evaluation	193
6.3 Comparing the results of the two approaches: quantitative evidence	196
6.3.1 Comparison of adoption decisions suggested by each approach	196
6.3.2 Methodological issues related to the outcome measures used in each approach	201
6.3.3 Summary of the quantitative comparison	206

6.4 Comparing the results of the two approaches: stakeholder survey	206
6.4.1 Survey methods.....	207
6.4.2 Survey results	208
6.4.3 Survey conclusions	211
6.5 Conclusions.....	212
Chapter Seven Conclusion.....	213
7.1 Introduction.....	213
7.2 Overview of thesis findings	213
7.3 What is the most appropriate economic evaluation approach to use when assessing genomic interventions?.....	215
7.3.1 How can appropriateness be assessed?	216
7.3.2 Which economic evaluation approach is most appropriate in this context?	217
7.4 Limitations of this research	221
7.5 Implications of this research	222
7.5.1 Contribution to health economics and other fields	223
7.5.2 Generalisability of these findings.....	224
7.5.3 Practical implications	224
7.5.4 Policy consequences	226
7.6 Future research	227
7.7 Concluding remarks.....	229
List of References	230
Appendix 1: Search strategies used to identify references which consider the impact of choice of economic evaluation approach on adoption decisions.....	257
Appendix 2: Consolidated Health Economic Evaluation Reporting Standards checklist and quality assessment of papers included in the literature review.....	259
Appendix 3: References included in the literature review which considers the impact of choice of economic evaluation approach on adoption decisions.....	287
Appendix 4: Summary of references reporting the results of both a welfarist and an extra-welfarist economic evaluation in the same study.....	289
Appendix 5: Search strategies used to identify references which summarise the methodological challenges faced when conducting economic evaluations of genomic interventions.....	309
Appendix 6: References included in the literature review which summarises the methodological challenges faced when conducting economic evaluations of genomic interventions.....	312

Appendix 7: Search strategies used to identify health economic evidence to support current genetic testing practice in CLL.....	315
Appendix 8: Genetic testing for chronic lymphocytic leukaemia in UK NHS laboratories	319
Appendix 9: Survey of haematologists	320
Appendix 10: List of questions in the haematologist survey	330
Appendix 11: Search strategies used to identify health economic evidence for the use of targeted next generation sequencing in all disease areas	336
Appendix 12: Summary of previous economic evaluations in chronic lymphocytic leukaemia conducted by the National Institute for Health and Care Excellence	339
Appendix 13: Descriptions of all model health states	342
Appendix 14: Transition probabilities used in the Markov model	344
Appendix 15: Resource use and overall costs for each health state	355
Appendix 16: Unit costs	369
Appendix 17: Microcosting study of chronic lymphocytic leukaemia genetic and genomic testing practice at the Oxford University Hospital NHS Trust	373
Appendix 18: Complication rates	382
Appendix 19: Costs associated with treating chemotherapy complications.....	384
Appendix 20: Costs associated with treating bone marrow transplant complications	385
Appendix 21: Average length of chemotherapy treatment	386
Appendix 22: Structured literature search conducted to identify utility values for chronic lymphocytic leukaemia patients in different disease states.....	387
Appendix 23: Parameter variations considered in univariate sensitivity analysis	391
Appendix 24: Sensitivity analysis results	401
Appendix 25: Scenario analysis results.....	412
Appendix 26: Cost-effectiveness planes	417
Appendix 27: Literature search for discrete choice experiments conducted in genetics or leukaemia	419
Appendix 28: Materials used in patient interviews.....	421

Appendix 29: Patient interview methods and results	432
Appendix 30: Further information on the pilot discrete choice experiment and the final experimental design	436
Appendix 31: Discrete choice experiment survey with risk graphics	438
Appendix 32: Discrete choice experiment survey without risk graphics	469
Appendix 33: Discrete choice experiment respondent opinions on genetic testing and attribute importance rankings	491
Appendix 34: Incorporating attribute interactions into Model A	492
Appendix 35: Sensitivity analysis for the cost-benefit analysis.....	496
Appendix 36: Extending the extra-welfarist economic evaluation to a societal perspective – methods and results	499
Appendix 37: Expected value of perfect information results presented graphically....	510
Appendix 38: Survey of healthcare decision-makers – methods and results.....	515
Appendix 39: Supporting materials provided to survey respondents	526
Appendix 40: Survey questions	530
Appendix 41: Survey of healthcare decision-makers – responses.....	537

Approximate number of words in this thesis: 50,000

List of tables

Table 2.1: Literature review summary statistics – impact of choice of economic evaluation approach on adoption decisions	26
Table 2.2: Summary of methodological issues arising in health economic evaluations of genomic technologies.....	36
Table 3.1: Variations in chronic lymphocytic leukaemia treatment pathways in other UK NHS Trusts	75
Table 4.1: A summary of the comparators considered in the Markov model	97
Table 4.2: Comparison between current genetic testing practice and future genomic testing practice	101
Table 4.3: Summary of the cost of spending a 28-day cycle in each model health state	110
Table 4.4: Utility weights used to calculate quality-adjusted life years in the model	112
Table 4.5: Base case results for the cost-effectiveness and cost-utility analyses.....	116
Table 4.6: Parameter variations which change the base case analysis results	119
Table 5.1: Discrete choice experiment attributes and levels	145
Table 5.2: Discrete choice experiment respondent characteristics	153
Table 5.3: Regression results and willingness-to-pay estimates.....	155
Table 5.4: Alternative specifications for Model A	158
Table 5.5: Willingness-to-pay for a hypothetical specification of genomic testing in chronic lymphocytic leukaemia	161
Table 5.6: Attribute levels for genetic and genomic testing in the cost-benefit analysis.....	170
Table 5.7: Main cost-benefit analysis results	173
Table 6.1: Results for the cost-effectiveness and cost-utility analyses from a societal perspective	185
Table 6.2: Results for the cost-effectiveness and cost-utility analyses from a societal perspective for patients aged 25.....	186
Table 6.3: Results for the cost-effectiveness and cost-utility analyses for a 20 year time horizon.....	188
Table 6.4: Results for the cost-effectiveness and cost-utility analyses for a ten year time horizon.....	188

Table 6.5: Results for the cost-effectiveness and cost-utility analyses for a one year time horizon.....	189
Table 6.6: Population expected value of perfect information	192
Table 6.7: ‘What if’ benefit transfer analysis results.....	195
Table 6.8: Summary of the adoption decisions suggested by each analytical approach..	197
Table 6.9: Summary of the decisions made by respondents in the survey	209
Table 7.1: Selecting the most appropriate economic evaluation approach to use when assessing genomic interventions – factors to consider.....	217

List of figures

Figure 2.1: The cost-effectiveness plane	17
Figure 2.2: Literature review results – impact of choice of economic evaluation approach on adoption decisions	25
Figure 2.3: Literature review results – methodological challenges faced when conducting economic evaluations of genomic interventions	35
Figure 3.1: Literature review results – health economic evidence on genetic testing practice in CLL.....	67
Figure 3.2: Clinical pathways for chronic lymphocytic leukaemia in the UK NHS.....	73
Figure 3.3: DNA sequencing timeline	78
Figure 3.4: Overview of targeted next generation sequencing process	80
Figure 3.5: Literature review results - economic evidence for the use of targeted next generation sequencing in all disease areas.....	83
Figure 4.1: A Markov model to simulate the disease pathway of ‘go-go’ patients	91
Figure 4.2: Comparator A pathways.....	93
Figure 4.3: Comparator B pathways.....	94
Figure 4.4: Comparator C pathways	95
Figure 4.5: Intervention One pathways.....	96
Figure 4.6: Intervention Two pathways.....	97
Figure 4.7: Cost-effectiveness acceptability curves for the cost-effectiveness analysis...	125
Figure 4.8: Cost-effectiveness acceptability curves for the cost-utility analysis.....	126
Figure 5.1: Example discrete choice experiment question both with and without risk graphics.....	147
Figure 5.2: Net benefit differences between genetic and genomic testing.....	175
Figure 5.3: Highest net benefit for either genetic or genomic testing.....	175

List of abbreviations

ACER	Average cost-effectiveness ratio
AIC	Akaike information criterion
Ass	Assumption
BMT	Bone marrow transplant
BIC	Bayesian information criterion
BMT	Bone marrow transplant
BR	Combination bendamustine and rituximab chemotherapy
BSC	Best supportive care
BWS	Best-worst scaling
CAIC	Consistent information criterion
Calc	Calculated / calculation
CBA	Cost-benefit analysis
CCA	Cost-consequence analysis
CDF	Cancer Drugs Fund
CEA	Cost-effectiveness analysis
CEAC(s)	Cost-effectiveness acceptability curve(s)
CI	Confidence interval
CLL	Chronic lymphocytic leukaemia
CLLSA	UK CLL Support Association
CMA	Cost-minimisation analysis
CMV	Cytomegalovirus
cnLOH	Copy neutral loss of heterozygosity
CNA	Copy number alteration
CR	Complete remission
CRD	Centre for Reviews and Dissemination
CT	Computerised tomography scan
CUA	Cost-utility analysis
CV	Contingent valuation
DALY(s)	Disability-adjusted life year(s)
DCE	Discrete choice experiment
DiSC	Differential staining cytotoxicity
DNA	Deoxyribonucleic acid
DOM	Dominated
EDTA	Ethylenediamine tetraacetic acid
ENBS	Expected net benefit of sampling
EO	Expert opinion
EQ-5D	EuroQol five dimensions questionnaire
ERG	Evidence review group

EU	European Union
EVSI	Expected value of sample information
EVPI	Expected value of perfect information
EXT.DOM	Extendedly dominated
FBC	Full blood count
FCA	Friction cost approach
FCR	Combination rituximab, cyclophosphamide and fludarabine chemotherapy
FISH	Fluorescent (or fluorescence) in situ hybridisation
g	Grams
GCSF	Granulocyte colony-stimulating factor
GNP	Gross national product
HCA	Human capital approach
HMDS	Haemato-molecular diagnostic laboratory
HTA	Health technology assessment
HUI	Health Utilities Index
ICER	Incremental cost-effectiveness ratio
IB	Ibrutinib
IgVH	Immunoglobulin heavy chain variable region
Int 1/2	Intervention One/Two
IV	Intravenous therapy
IVIG	IV immunoglobulin treatment
Kb(s)	Kilo base pairs (1,000 bp)
LFT	Liver function test
LY(s)	Life year(s)
m	Metres
Mb	Mega base pairs (1,000,000 bp)
mcg	Micrograms
MeSH	Medical Subject Headings
mg	Milligrams
ml	Millilitres
MLPA	Multiplex ligation-dependent amplification
MNL	Multinomial logit regression
MRD	Minimal residual disease
MTO	Medical technical officer
MXL	Mixed logit regression
NGS	Next generation sequencing
NICE	National Institute for Health and Care Excellence
NHS	National Health Service
NHS EED	The NHS economic evaluation database
Ofa (or O)	Ofatumumab
OOP	Out of pocket

OUH	Oxford University Hospitals
P	Probability
PCR	Polymerase chain reaction
PFS	Progression-free survival
PR	Partial remission
PSA	Probabilistic sensitivity analysis
QALY(s)	Quality-adjusted life year(s)
R	Residual probability
RCT	Randomised controlled trial
RR	Response rate
RUT	Random utility theory
SD	Standard deviation
SG	Standard gamble
Soc	Societal
SOP	Standard operating procedure
TP	Transition probability
UK	United Kingdom
US(A)	United States of America
U&E	Urea and electrolytes test
VAF	Variant allele frequency
VAS	Visual analogue scale
VOI	Value of information
WES	Whole-exome sequencing
WGS	Whole-genome sequencing
WTP	Willingness-to-pay (or willing-to-pay)

Chapter One | Introduction

1.1 Rationale for the thesis

This thesis explores a number of empirical and methodological issues that arise when conducting economic analyses of genomic diagnostic technologies for multifactorial genetic diseases in the UK National Health Service (NHS). Successfully addressing these issues is important from a health economics perspective as this could facilitate the more efficient allocation of scarce healthcare resources. The aim of this chapter is to explain in more detail why this is an important topic of research in health economics. A useful starting point is to describe the clinical utility of genomic information in health care and to consider how tests which generate this information have evolved in the past decade. The challenges associated with conducting economic evaluations of these tests are then outlined, which in turn provides the rationale for this thesis.

Deoxyribonucleic acid (DNA) is a nucleic acid that carries all of the biological information required to build and maintain an organism. DNA is found in cells and organised mostly into structures called chromosomes. DNA is made up of a sequence of chemical bases and contains thousands of defined segments called genes, which influence the characteristics of an organism. Genes can sometimes acquire DNA sequence changes (mutations) which are either inherited (germline mutations) or arise by chance (somatic mutations). This thesis focuses on somatic mutations, which can have a variety of consequences. Some are benign and have no effect on biochemical processes in the body; others have minor effects

that will never become visible. However, some mutations can have more obvious effects that either result in disease, alter disease risk or modify response to treatment.

Genetic tests are assays that are targeted to specific genes of interest, or can identify large chromosomal changes. These tests require DNA, which is often taken from blood or soft tissue samples (or bodily fluids such as urine, in the case of infectious disease). The information generated by these tests can be diagnostic (e.g. identifying a chromosomal condition), prognostic (e.g. indicating the likely course of disease) or inform clinical management (e.g. predicting likely response to therapy). The quantity and range of genetic tests available to patients has increased significantly over the last 20 years, and genetic testing is now established practice in many disease areas in developed countries (e.g. *BRCA1/2* testing in breast cancer [1]). However, the most commonly used tests (e.g. karyotyping, which requires scientists to visually inspect chromosomes using a microscope) are not sufficiently reliable or lack the resolution to provide detailed information on multiple genetic changes across the whole genome. There is increasing evidence that such information could enable improved disease stratification and individually tailored therapies, which could benefit patients (by reducing morbidity and mortality) and the NHS (by reducing the prescription of expensive medicines) [2,3].

Attention is now turning towards genomic interventions with the potential to meet these objectives, such as next generation sequencing technologies. These technologies can simultaneously scrutinise multiple genes and their interrelationships in order to identify their combined influence, and have broad-ranging application in medicine, showing promise in allowing disease management to be stratified. These tests may be particularly

informative in multifactorial genetic diseases (e.g. inflammatory bowel disease [4]), where individual gene mutations contribute, but are insufficient alone, to cause the disease [5].

Despite some success stories, genomic technologies have had a limited impact on clinical practice to date. Only 1-2% of all drugs listed by the US Food and Drug Administration contain pharmacogenomic information on their label [6] and testing is strongly recommended in only four cases (including the use of cetuximab to treat EGFR-expressing colorectal cancer, and the use of trastuzumab to treat HER2-positive breast cancer) [7]. Internationally, adoption rates for genomic technologies vary considerably. For example, 90% of insured individuals in the USA have an insurance plan which covers Oncotype DX (a diagnostic test for breast cancer) [8], but the UK National Institute for Health and Care Excellence (NICE) has only recently provisionally approved the use of this test [9].

Differences in the uptake of these technologies have arisen, in part, due to a lack of clinical evidence (on test effectiveness) and ethical concerns regarding the treatment of secondary findings. However, translational research evidence is also scarce [10], particularly economic evidence [11]. In the UK NHS, such evidence is typically generated by conducting an economic evaluation, defined as *“the comparative analysis of alternative courses of action in terms of both their costs and consequences”* [12]. Economic evaluations are currently undertaken in many disease areas and provide information which can be used to guide the efficient allocation of scarce healthcare resources. The usual approach, as recommended by NICE [13], is to conduct a form of extra-welfarist economic evaluation, in which outcomes are measured in health-related terms. Options include cost-effectiveness analysis (CEA), in which outcomes are measured in natural units such as life-

years, and cost-utility analysis (CUA), with outcomes measured using quality-adjusted life years (QALYs) which combine information on longevity and quality of life. An alternative approach which is infrequently used at present in the context of health care is to conduct a welfarist economic evaluation such as a cost-benefit analysis (CBA), in which outcomes are measured in monetary terms. The aims of these two approaches differ: extra-welfarist approaches aim to maximise health outcomes [14,15], whereas the welfarist approach is broader, and can capture both health and non-health outcomes (e.g. process-related outcomes such as improvements in waiting time) [15-17].

To date, most economic evaluations of genetic or genomic technologies have taken an extra-welfarist approach. However, health economists have begun to question whether this is always appropriate in genomics [18]. There is a particular debate surrounding the need to capture information on non-health outcomes (e.g. the value to patients of information from test results), which might not be captured by QALYs. Other challenges surround the analytical approach (e.g. should a health service or societal perspective be used?) and how to incorporate genomic test effectiveness data into economic analyses when the evidence base linking this data to health outcomes is generally underdeveloped.

Health economists are uncertain as to whether existing approaches to economic evaluation are sufficient to evaluate all genomic technologies. Clarification is required because different approaches may lead to different decisions regarding the adoption of these technologies in routine clinical practice. More advanced (and expensive) genomic tests are starting to emerge (e.g. whole genome sequencing), providing further challenges for existing methods. If agreement could be reached on the most appropriate methods of

economic assessment in this context, health economists would be better placed to contribute to translational research in genomics [19].

This thesis explores a number of these methodological challenges. The main focus is on identifying the most appropriate economic evaluation approach to use when assessing genomic interventions, and the advantages and disadvantages of different approaches are demonstrated using a clinical case study. The following section introduces this case study: the use of targeted next generation sequencing to guide treatment decisions in chronic lymphocytic leukaemia (CLL). The remaining sections of this chapter outline the thesis aims and objectives, describe the structure of the thesis and explain how this work will contribute to the literature, both in health economics and beyond.

1.2 Clinical case study

CLL is the most common adult leukaemia [20-23]. It occurs mainly in individuals over the age of 60, is incurable, and is characterised by a chronic relapsing course, requiring regular treatment and thus representing a significant burden for patients, their families and health services. Morbidity effects arising from chemotherapy treatment are notable: two-thirds of patients die from treatment complications [24]. Genetic abnormalities can provide important information about likely response to chemotherapy. However, current genetic tests (such as fluorescence in situ hybridisation testing) have limited resolution to detect these mutations. Genomic tests (such as targeted next generation sequencing) could improve treatment stratification by detecting more of these mutations, as well as additional changes such as genomic complexity (described in Chapter Three). Such tests could therefore enable CLL treatments to be targeted more effectively, reducing morbidity

and mortality, improving patient well-being and lowering the cost burden of CLL for healthcare providers. This cost burden is a particular concern in CLL because alternative treatments for patients identified as likely non-responders include drugs such as ibrutinib, which is expensive. Genomic information is also used to inform the stratification of patients to expensive therapies in other cancers, including breast and colorectal cancer. This makes CLL an excellent case study for the exploration of methods within this thesis. Chapter Three provides a more detailed description of both CLL and targeted next generation sequencing.

1.3 Thesis aims and objectives

This thesis aims to address the following research question: what is the most appropriate economic evaluation approach to use when assessing genomic interventions? The primary hypothesis of this thesis is that the welfarist approach to economic evaluation is most appropriate in this context (rather than an extra-welfarist approach) as it is better placed to capture information on the non-health outcomes that are important to patients undergoing genomic testing. This hypothesis will be tested by conducting a number of analyses using data from a study which has developed a genomic test to guide treatment decisions in CLL [24]. Four key objectives underlie this research question:

1. Summarise the methodological challenges faced when conducting economic evaluations of genomic technologies;
2. Evaluate the existing clinical and economic evidence for the use of genetic and genomic testing to guide treatment decisions in CLL;
3. Estimate the costs and outcomes associated with genetic and genomic testing in CLL from alternative perspectives;

4. Use these data to conduct welfarist and extra-welfarist economic evaluations of the use of genetic and genomic testing in CLL, and explore the implications of different approaches for decision-making.

1.4 Structure of the thesis

This thesis has seven chapters. This section provides a synopsis of each of these chapters.

Chapter Two

Chapter Two provides the methodological background to the thesis by introducing the process of economic evaluation and describing the welfarist and extra-welfarist approaches in detail. A literature review is then presented which evaluates the extent to which choice of approach can impact on the adoption decisions recommended by economic evaluation studies. The chapter then focuses more specifically on conducting economic evaluations of genomic technologies, and a second literature review summarises the methodological challenges associated with conducting economic evaluations in this context. The chapter concludes by describing how the results of these reviews will be taken forward to address the main research questions in this thesis.

Chapter Three

This chapter provides information on the clinical case study which is used to demonstrate the advantages and disadvantage of different economic evaluation approaches. The chapter begins by describing current patient management and treatment practice for CLL patients in the UK. Particular consideration is given to the rationale for using genomic testing to guide treatment decisions, and evidence is presented on current and likely future

use of genetic and genomic tests, collected via a survey of UK haematologists. Background information on targeted next generation sequencing is also provided, and the advantages of using this test in CLL are described. Finally, the economic evidence for the use of targeted next generation sequencing in all disease areas is summarised.

Chapter Four

Chapter Four describes the first of two analyses conducted to establish the most appropriate economic evaluation approach to use when assessing genomic interventions. The analysis presented in this chapter is an extra-welfarist economic evaluation of genomic testing in CLL which takes two forms: a CEA, with outcomes measured using life-years, and a CUA, with outcomes measured using QALYs. In both cases, five strategies are compared reflecting different combinations of genetic or genomic testing (or no testing), and the addition of new therapies such as ibrutinib to current or future treatment pathways (and at different stages in these pathways).

Chapter Five

Chapter Five describes the second analysis conducted to establish the most appropriate economic evaluation approach to use when assessing genomic interventions: a welfarist economic evaluation (a CBA) which evaluates the costs and consequences of using genetic and genomic testing to guide chemotherapy treatment decisions for CLL patients. This CBA required data on willingness-to-pay (WTP) for testing, which in this instance necessitated the use of a discrete choice experiment (DCE) to collect such data. Although DCEs are a potentially informative source of data on WTP for genomic testing, operationalising this approach in a genomic context is not straightforward, and there are several

methodological obstacles preventing the more widespread use of this technique. This chapter also addresses some of these issues. In particular, a comparison of visual and non-visual approaches to presenting genomic risk information to patients is described, and the impact of different DCE design decisions on the CBA results is considered.

Chapter Six

Chapter Six presents a comparison of the welfarist and extra-welfarist economic evaluations described in Chapters Four and Five. As these approaches ask fundamentally different resource allocation questions, several analytical extensions were required to facilitate such a comparison. A description of these extensions is provided, and this additional evidence then informs a quantitative comparison of the results of the two approaches. The chapter concludes by presenting the results of a stakeholder survey which assesses the degree to which the results of the different economic evaluation approaches capture the true costs and benefits of genomic testing in CLL.

Chapter Seven

Chapter Seven concludes the thesis by considering the evidence presented in Chapters Four, Five and Six and making some recommendations regarding the most appropriate economic evaluation approach to use when assessing genomic interventions. The generalisability, practical implications and policy consequences of the results are also considered. Finally, issues that have arisen during the preparation of this thesis are discussed and areas for future research are outlined.

1.5 Contribution to the literature

The aim of this thesis is to contribute to three fields of research: health economic evaluation methodology, DCE methodology and the health economics of CLL patient care.

The first contribution to the economic evaluation methodology literature is made by the literature reviews which evaluate the impact of choice of approach on adoption decisions and consider the challenges associated with conducting economic evaluations in genomics.

The key result from the first of these literature reviews is that for every four studies applying both the welfarist and extra-welfarist approaches, one shows limited or no concordance between the results of the different forms of economic evaluation, suggesting that choice of approach can impact on adoption recommendations. The key result from the second of these literature reviews is that economic evaluations of genomic technologies present a particular challenge for health economists due to the accumulation of a variety of methodological challenges in this context. Further contributions to the economic evaluation methodology literature are made by the comparison of economic evaluation approaches and the survey which assesses the degree to which different approaches capture the true costs and benefits of genomic testing in CLL. The key finding from these analyses is that the extra-welfarist approach to economic evaluation is most appropriate when assessing genomic interventions. Finally, to date, only three welfarist economic evaluations have been conducted based on a DCE [25,26]; the CBA presented in Chapter Five therefore also advances this literature.

The second field of research to which this thesis contributes concerns DCE methodology. Few DCEs have been conducted in genetics, even fewer have considered patient preferences for genomic technologies, and only one study has conducted a comparison of

visual and non-visual approaches for the presentation of risk information [27]. Further background on the use of DCEs in this context, as well as a more detailed description of the risk presentation literature is provided in Chapter Five. The main finding from these analyses is that estimates of WTP differ when risk presentation is varied, suggesting that visual and non-visual approaches to risk presentation may measure different cognitive processes.

Finally, the health economic evidence base in CLL is extremely limited. There are few economic evaluations of any genetic or genomic tests, limited data exists on current genetic testing practice in the UK NHS and there is no information available on likely future genomic testing practice in this context. Furthermore, no economic evaluations of the use of ibrutinib have been conducted. This thesis contributes to the CLL literature on all of these counts. The key finding from these analyses is that the use of genomic testing to guide treatment decisions in CLL is unlikely to represent an efficient use of scarce healthcare resources in the UK NHS.

1.6 Conclusion

This chapter has provided an overall introduction to the thesis, describing the rationale, objectives and contribution that this thesis will make to various literatures. Chapter Two will now describe the methodological background to this work.

Chapter Two | Conducting economic evaluations of genomic technologies: methodological challenges

2.1 Introduction

This chapter begins by introducing the discipline of economics and explaining how economic methods are applied to health and health care. An area of research within health economics – economic evaluation – is then described, and two approaches to economic evaluation are presented: the welfarist and the extra-welfarist approach. These approaches differ in terms of their aims and practical application, and the most commonly applied approach has varied over time. Section 2.3 considers how the debate concerning the most appropriate approach has evolved, and a literature review is presented which evaluates the extent to which choice of approach can impact on the adoption decisions recommended by economic evaluation studies. Section 2.4 focuses more specifically on conducting economic evaluations of genomic technologies. A second literature review summarises the methodological challenges associated with conducting economic evaluations in this context, identifying issues where there is consensus over potential solutions and those where there is uncertainty and further research is required. The chapter concludes by describing how the results of these literature reviews will be taken forward to address the main research question posed in this thesis: what is the most appropriate economic evaluation approach to use in genomics?

2.2 Background to economics and health economic evaluation

2.2.1 Rationale for economic evaluation

Economics is the study of how individuals and societies choose to allocate scarce resources (e.g. land, labour, capital, time) among competing alternative uses (e.g. the production of different goods and services). Given that resources are limited, it is important that they are used as efficiently as possible to maximise the economic benefits (welfare gains) to society. Standard economic theory suggests that the best way to achieve this goal is to allow consumers and producers to interact in a free market, using price as a rationing tool to equalise demand and supply. However, the free market approach requires several assumptions to hold, including the existence of perfect knowledge and certainty (*“consumers know exactly what they want, when they want it, and when they can get it”* [28]), the absence of externalities (consequences of economic activity that are experienced by unrelated third parties [28]), and the existence of many producers with no market power. In some sectors of the economy, these requirements are more easily met (e.g. markets for consumer goods such as clothes and food) but in others (e.g. national defence services) it is not as straightforward. If these assumptions do not hold in a particular sector of the economy, the free market may not be the most efficient way to allocate scarce resources, and alternative rationing tools are required.

Allocating scarce resources in health care is particularly challenging because the free market fails in several ways. First, it is difficult for individuals to predict when they will require health care. In a free market, firms will respond to this uncertainty by offering insurance. However, healthcare insurance markets suffer from two fundamental problems: moral hazard (individuals possessing insurance may change their behaviour) and adverse

selection (individuals know more about their risk status than those offering insurance). Second, externalities exist (e.g. vaccinations benefit both participating individuals and also the rest of society). Third, there is an imperfect principal agent relationship: clinicians act on behalf of patients to both demand and provide health care, which could potentially lead to more health care being delivered than is necessary. Finally, professional licensure is required as mistakes in health care can have serious consequences. However, this gives clinicians a degree of market power, which could be exploited to reduce the supply of health care. Taken together, these problems can prevent the free market from reaching equilibrium in health care and thus allocating healthcare resources efficiently. As a consequence, it is generally accepted that government intervention is required to allocate health care resources efficiently [28].

The extent to which governments intervene to allocate resources in health care varies across countries, and depends on the aims of a healthcare system. The UK National Health Service (NHS) is generally said to have three basic aspirations: to provide “comprehensive, high quality medical care [...] to all citizens on the basis of professionally judged medical need without financial barriers to access” [29]. However, as healthcare resources are limited, these aims are said to be an “*inconsistent triad*” – it is not possible to achieve all three simultaneously and stay within a budget [29]. The core principles of the UK NHS were stated by Aneurin Bevan in 1948 and included commitments that health care will be available to all and free at the point of delivery [30]. As a consequence of this universal coverage, the NHS is considered by some to compromise on the range and quality of services offered to patients [29]. This can mean that the most clinically effective interventions are not always offered, or are only offered to patient subgroups. Healthcare

resource allocation decisions such as these can be made by using waiting lists, or on a case by case basis, but these approaches could lead to random, implicit or subjective decisions. An alternative approach is to make these decisions using evidence from systematic and explicit decision tools: economic evaluation is such a tool.

2.2.2 Types of economic evaluation

Economic evaluation is commonly defined as *“the comparative analysis of alternative courses of action in terms of both their costs and consequences”* [12], and it can be used to compute and compare the costs and outcomes of different healthcare interventions with a view to identifying those which represent the best value for money. Economic evaluations of healthcare interventions commonly employ either a ‘welfarist’ or ‘extra-welfarist’ approach. The aim of the welfarist approach is considered to be to maximise societal welfare with reference to a societal budget constraint [14,15], while the aim of the extra-welfarist approach is generally believed to be to maximise health effects in a resource-constrained health system [15-17]. So broadly speaking, extra-welfarism considers resource allocation issues following budget allocation (e.g. should more of the NHS budget be spent on drugs for diabetes or cancer care?), whereas welfarism provides information which can be used to inform the efficient determination of budgets in the first place (e.g. should more money be allocated to the NHS budget or the defence budget?) [31]. There is a debate within health economics on the extent to which the two approaches are linked and the evolution of this debate is described in Section 2.3.1. Practically speaking, both approaches measure healthcare costs in monetary terms (using the same methods), but each differs in the measurement of health outcomes. The following subsections describe these differences.

The extra-welfarist approach to economic evaluation

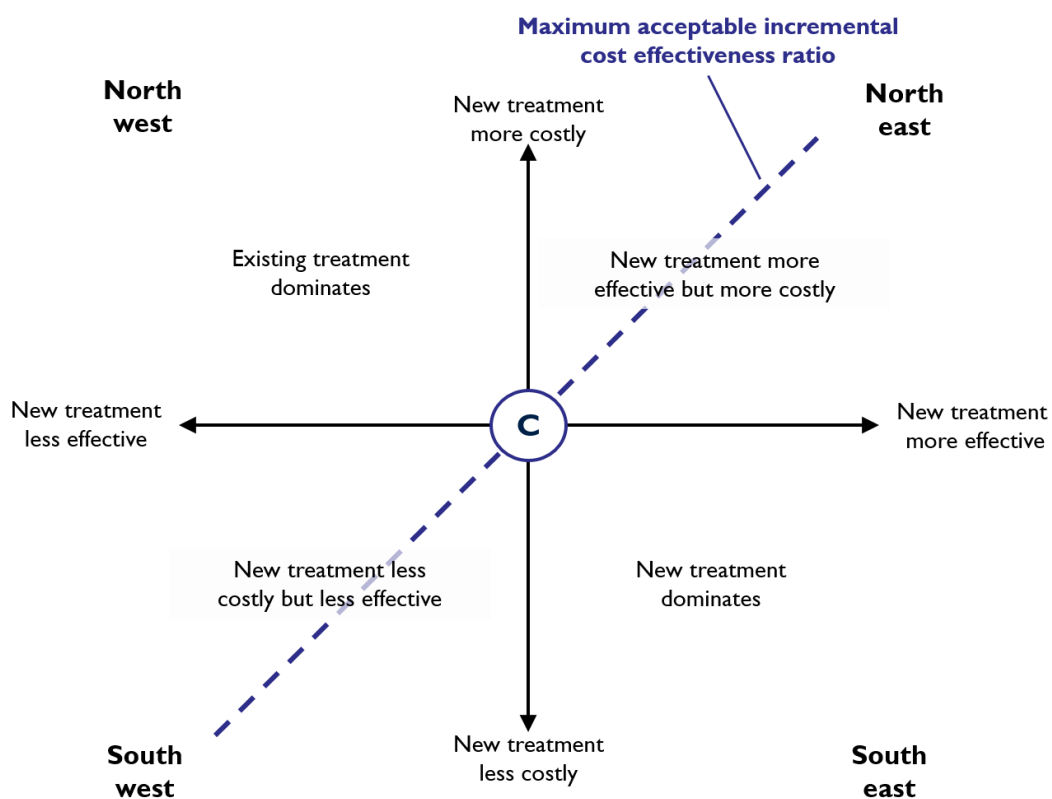
Health economists distinguish between several ways of operationalising the extra-welfarist approach. Cost-minimisation analysis (CMA) used to be recommended in situations where two or more alternatives under consideration achieved a given health outcome to the same extent [12], with the preferred intervention being the least costly option. However, although CMAs are still being published [32] (and are accepted by HTA agencies in evidence submissions for biosimilars when evidence exists of equivalent effectiveness between a biosimilar and a comparator [33,34]), CMAs are no longer considered to be a form of full economic evaluation [35]. Indeed, it has recently been noted that *“the remit of CMA in trial-based economic evaluation is [...] even narrower than previously thought, suggesting that CMA is not only dead but should also be buried”* [32].

The two principal ways in which the extra-welfarist approach is operationalised are cost-effectiveness analysis (CEA) and cost-utility analysis (CUA). Both measure outcomes in health-related terms: CEA uses natural units such as cases detected (e.g. in a screening programme) or number of life-years gained from using an intervention, while CUA (effectively a variant of CEA) uses quality-adjusted life years (QALYs), which combine information on longevity and quality of life (disability-adjusted life years are also used in some settings). In both cases, when an economic evaluation is conducted, the cost and outcome differences between a new intervention and its comparator are calculated [12]. The decision rule used when deciding which intervention to adopt is straightforward when a new intervention is both less costly and more effective than its comparator (the new intervention is said to dominate), or when the new intervention is more costly and less effective (the comparator dominates). When a new intervention is both more costly and

more effective than its comparator, the cost and outcome differences are expressed as a ratio (the incremental cost-effectiveness ratio, or ICER). This ICER is then compared to a threshold (the maximum acceptable ICER) which represents the maximum amount that society is willing to pay for an additional unit of health outcome. If the ICER falls below this threshold, the new intervention is said to be cost-effective, and is likely to be adopted [12].

Incremental costs and outcomes can be expressed visually using the cost-effectiveness plane (Figure 2.1) [36]. This consists of four quadrants, each reflecting different combinations of positive and negative incremental costs and outcomes. The most common scenario is that the ICER for a new intervention versus its comparator (C) falls in the north east quadrant, with the adoption decision depending on the maximum acceptable ICER.

Figure 2.1: The cost-effectiveness plane



The welfarist approach to economic evaluation

The application of the welfarist approach to economic evaluation usually takes the form of cost-benefit analysis (CBA), with health outcomes valued in monetary terms using methods such as the human capital or revealed preference approach [12]. These approaches are described in more detail in Chapter Five. Results are usually expressed as either a ratio of benefits to costs, or as a simple sum representing the net benefit of one intervention compared to another. The decision rule used when deciding which intervention to adopt is therefore straightforward: if a new intervention generates a positive net benefit (or a ratio of benefits to costs greater than one), it is worthwhile to adopt this intervention. The use of a common unit (money) to value outcomes means that healthcare interventions in a wide variety of disease areas can be easily compared in terms of their net benefit.

2.3 Welfarism versus extra-welfarism: which approach should be used?

2.3.1 Evolution of the debate over time

Section 2.2 described how the welfarist and extra-welfarist approaches to economic evaluation differ in terms of their aims and practical application. One of the consequences of these differences is that the most commonly applied approach has varied over the past 50 years. Although the welfarist approach initially dominated health economics during the 1960s and 1970s, the limitations of this approach became apparent over time. These included distributional concerns (CBA may favour diseases of the affluent over those of the poor [14,16,37]), and a lack of confidence in the use of the human capital approach (given that this approach may discriminate against those outside the labour force [38-41]). The consensus therefore evolved during the 1980s to favour the extra-welfarist approach, considered at that time to be more straightforward to operationalise in healthcare. An

increased demand for economic evidence, a rise in the number of health technology assessment (HTA) agencies, and the development of key metrics (e.g. the QALY [42]) and 'reference cases' for economic evaluation all contributed to this shift. As a result, the extra-welfarist approach grew to dominate decision-making in healthcare and was subsequently adopted by most HTA agencies around the world.

Health economists returned to this debate in the 1990s as there remained some dissatisfaction with the narrow health-oriented goals of extra-welfarism. It was recognised that this approach did not always provide useful information about how to allocate scarce resources [43], and it was noted that the welfarist approach may be more appropriate in specific decision contexts (e.g. when budgets are variable, use of the welfarist approach avoids spending energy and research time identifying a precise QALY value) [44]. In addition, distributional issues also impacted on the extra-welfarist approach: life expectancy is systematically related to income, wage rates and incomes are closely linked, and the valuation of costs by reference to market prices is also connected to income distribution (willingness-to-pay (WTP) - linked to ability to pay - drives demand, which, when equalised with supply, sets prices) [37]. Furthermore, there were practical challenges associated with operationalising extra-welfarism, including a reliance on the use of rules of thumb in decision-making [45-48], and the difficulty of identifying a single healthcare budget to which an extra-welfarist economic evaluation should refer [17]. These issues prompted a renewed interest in methods to monetise health benefits within a welfarist framework, such as the contingent valuation and discrete choice experiment (DCE) approaches [49-52].

Achieving a consensus on the most appropriate approach would become less important if it could be shown that the two approaches are equivalent. Historically, the two approaches were believed to be linked: Mishan stated in 1971 that *“To be rather rude about it, the analysis of cost-effectiveness can be described as a truncated form of cost-benefit analysis”* [53]. In the 1990s, it was argued that CEAs and CUAs could be converted into CBAs [17,39,54,55], and that both approaches would always yield the same answer to a given decision problem, providing certain assumptions held (e.g. the marginal value of a QALY is the same in both approaches [17,39,55-57]). This consensus has now shifted and the two approaches are believed to ask fundamentally different allocation questions: extra-welfarism aims to determine the least costly way to achieve a goal, while one of the aims of welfarism is to determine whether a goal is worth achieving [14,58]. Also, the assumptions that underlie the potential equivalence of the two approaches are now viewed as unlikely. For example, few believe that a single societal value for the WTP for a QALY exists [14,39,41,59]; instead, WTP is believed to increase with higher wealth and life expectancy, and decrease with improved health status [14,16,39,41,44,56,58-60].

Given that equivalence is unlikely, and the two approaches may give different answers to the same question [17,39,61]), health economists are beginning to reconsider the relative merits of the different approaches in specific decision contexts. The increasing number of complex interventions poses a particular problem [62]. Such interventions can involve several interacting components, numerous and difficult behaviours required by those delivering (or receiving) the intervention, and numerous outcomes [63]. Examples include interventions to control the behaviour of at risk individuals (e.g. drug users) and interventions based around interlinked packages of care (e.g. clinical genetics services)

[62,64]. Genomic technologies have a number of similar characteristics. In addition to their direct impact on health, these interventions may provide a broader range of benefits including satisfaction from the process of care or the acquisition of health-related information, and improved capabilities. Some of these benefits fall outside the scope of CUA or CEA, hence these interventions do not fit neatly into the economic evaluation framework applied by the major HTA agencies.

To date, few concessions have been made to accommodate complex interventions within current analytical frameworks. Some have suggested that the welfarist approach may better capture their benefits [41,57], but CBAs are rarely undertaken in health care. At least two HTA agencies (the Canadian Agency for Drugs and Technologies in Health and the Pharmaceutical Benefits Advisory Committee in Australia) permit CBAs to be submitted as supplementary analyses, but such analyses rarely influence HTA decisions. Even in scenarios where it has been recognised that an alternative regulatory context may be appropriate (e.g. HTAs conducted within the National Institute for Health and Care Excellence (NICE) Diagnostics Assessment Programme), the analytical approach follows standard practice, with benefits measured using QALYs [65,66]. One exception is the evaluation of public health interventions by NICE: CBAs are permitted in this context if extra-welfarist approaches are judged to be unsuitable [67]. However, this is an outlier; HTA methods are generally not adapted to reflect the characteristics of specific subgroups of healthcare interventions.

2.3.2 Can the choice of economic evaluation approach impact on the adoption decisions recommended by economic evaluation studies?

Given an increasing awareness that healthcare interventions may yield non-health benefits, and that the welfarist and extra-welfarist approaches to economic evaluation are unlikely to be equivalent, it is increasingly important that the most appropriate economic evaluation approach is used to enable decision-makers to be confident in their adoption decisions. To place this analytical decision into context, this section presents a literature review which evaluates the extent to which choice of approach can impact on the adoption decisions recommended by economic evaluation studies.

Literature review aims, objectives and methods

A structured literature review was conducted to identify references which report the results of both a welfarist and extra-welfarist economic evaluation in any disease area in the same study. For the purposes of this review, CEA, CUA and CMA were considered to be extra-welfarist approaches. Cost-consequence analysis (CCA) – in which a variety of costs and outcomes are presented in a disaggregated manner – was also considered to be an extra-welfarist approach. Partial economic evaluations (e.g. cost analyses) were not included. CBAs were only included if health outcomes were valued in monetary terms (e.g. using WTP, or by estimating the value of a life lost).

Literature searches were undertaken during August 2013 in Embase, Medline, Econlit, the UK National Health Service Economic Evaluation Database (NHS EED), the UK NHS HTA programme, and The Cochrane Library. Search strategies combined keyword variations using the logical operator 'OR', maximising sensitivity and minimising false negatives.

Keywords were identified by examining relevant references in the literature, previous related literature reviews, and Medical Subject Headings (MeSH). Searches were limited to full length references published in English (to enable in depth reviews to be carried out), and references published from 1990 onwards (to ensure that the results were applicable to the present setting). An exploratory search for earlier references was undertaken in Embase prior to undertaking the main search. However, no relevant references were identified. Details of the search strategies are provided in **Appendix 1**.

Information was extracted from references which met the inclusion criteria on the types of economic evaluation conducted, whether justification for adopting specific approaches was provided, the study results, whether the most appropriate approach was identified, and source of funding. A judgment was also made as to whether the results showed full concordance (all forms of economic evaluation suggested the same adoption decision), limited concordance (weak evidence of concordance e.g. only two out of three forms of economic evaluation supported a particular adoption decision, with limited confidence), or no concordance (each approach suggested a different adoption decision, often with little confidence in the results). The quality of the methods used in each study was assessed using the Consolidated Health Economic Evaluation Reporting Standards checklist [68] (provided in **Appendix 2**). For each item on the checklist, studies were assigned a score of one if it was fully addressed, 0.5 if it was partially addressed, zero if it was not addressed, or no score if the item was not applicable (N/A). All items were weighted equally and studies were classified as low quality (fulfilled <50% of applicable quality criteria), average quality (fulfilled between 50% and 80% of applicable criteria), or high quality (fulfilled >80% of applicable criteria) [69].

Literature review results

Figure 2.2 summarises the literature search results. 1737 references were identified, with 1035 titles remaining after removing duplicates. Of these, 15 met the inclusion criteria. ten references were added from other sources (e.g. reviewing the references of included studies, searching the publication records of key authors in this field), so 25 references in total were included in the review (**Appendix 3** presents a complete list). These references were published between 1992 and 2012, with 17 references published after 2000. The main reason for exclusion was that a full CBA had not been performed (with health outcomes valued in monetary terms); instead these excluded studies often reported a simple cost comparison.

The material extracted from each reference is detailed in **Appendix 4** and summarised in **Table 2.1**. The references spanned several clinical areas, with only three areas covered by multiple references (HIV-related interventions, vaccination, and back pain management). Overall, 13 studies (52%) considered public health interventions. The most common study locations were the USA (36% of studies), the UK (20%) and Canada (16%). There was no pattern to the comparisons made in these studies (comparators included 'do nothing', 'current practice', 'placebo' and other treatment options). Most (80%) of the studies were not commercially funded. The only agencies that funded more than one study were the US Centers for Disease Control and Prevention (four studies: 16%), the UK Medical Research Council (two studies: 8%) and the UK HTA Programme (two studies: 8%).

Figure 2.2: Literature review results – impact of choice of economic evaluation approach on adoption decisions

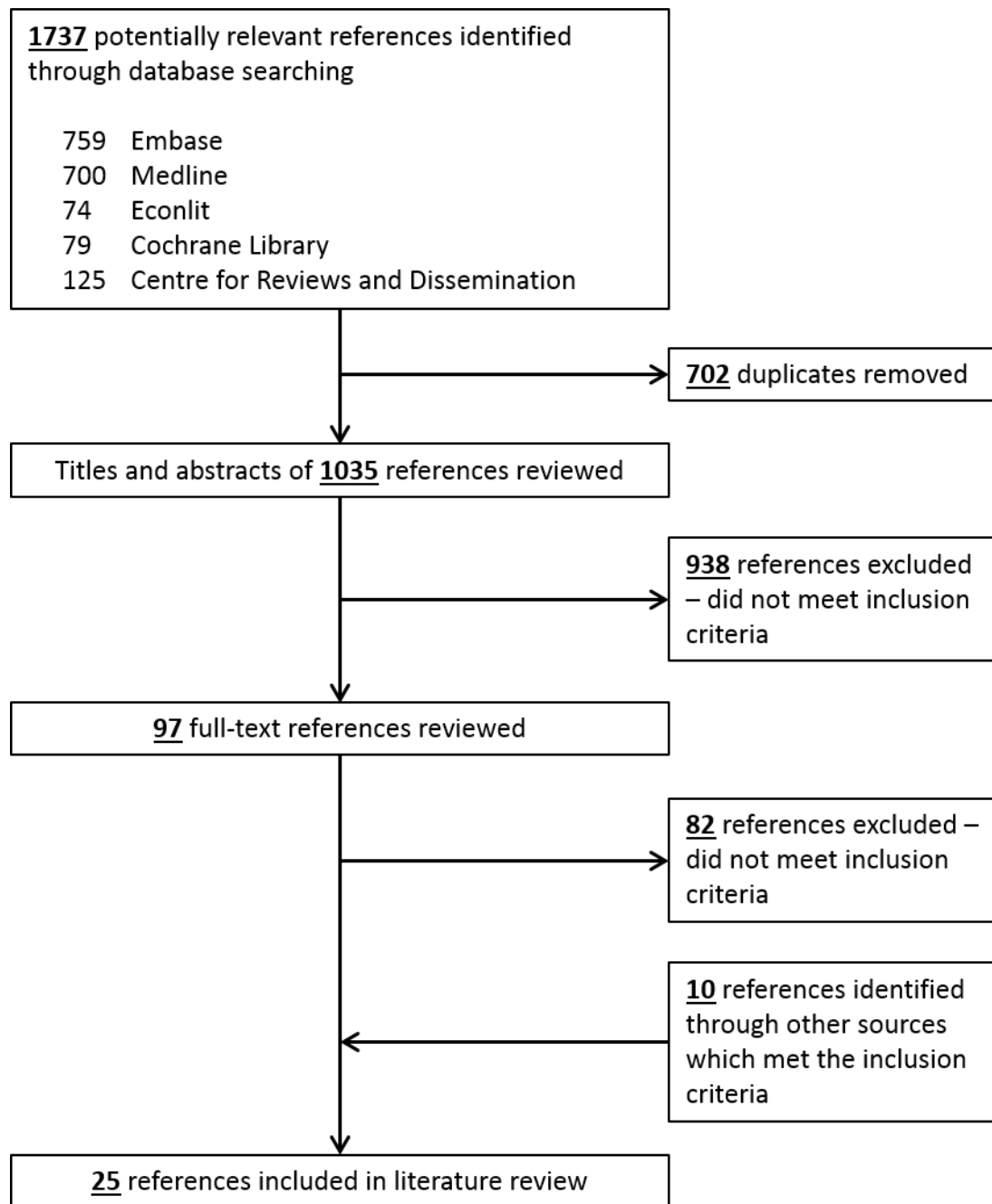


Table 2.1: Literature review summary statistics – impact of choice of economic evaluation approach on adoption decisions

Statistic	Category	Number of studies in each category (%)
Source of funding	Commercial	2 (8%)
	Not commercial	20 (80%)
	No information given	3 (12%)
Intervention context	Vaccination	3 (12%)
	HIV-related interventions	3 (12%)
	Managing back pain	2 (8%)
	17 other clinical areas ^a	1 (4%)
Study setting	USA	9 (36%)
	UK	5 (20%)
	Canada	4 (16%)
	Netherlands	2 (8%)
	Sweden	2 (8%)
	Israel / India / Iran	1 (4%)
Type of extra-welfarist economic evaluation conducted	CEA alone	11 (44%)
	CUA alone	6 (24%)
	CEA and CUA	8 (32%)
Method used to estimate utilities within CUA	EQ-5D	6 (43%)
	HUI	3 (21%)
	SG / VAS	2 (14%)
	DALYs	2 (14%)
	Alternative multi-attribute utility instrument	1 (7%)
Justification provided for welfarist approach?	Yes	5 (20%)
	No	20 (80%)
Concordance between results?	Yes	18 (72%)
	Limited	3 (12%)
	No	4 (16%)
Most informative economic evaluation results identified?	Yes	3 (12%)
	No	22 (88%)

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DALY = disability-adjusted life years; DCE = discrete choice experiment; EQ-5D = EuroQol five dimensions questionnaire; HUI = Health Utilities Index questionnaire; SG = standard gamble; VAS = visual analogue scale; WTP = willingness-to-pay. ^a Includes addiction treatment, promoting physical activity, prenatal screening, psoriasis, migraine, hepatitis A, TB, folic acid fortification, hypertension, health impact of smoke alarms, heart disease, antiemetic treatment, bronchitis, mental health, neurosurgery, surveillance for ocular hypertension, hernia surgery.

All studies reported the results of a CBA. None reported the results of a CMA or CCA. Eleven studies (44%) reported the results of a CEA alongside a CBA, six studies (24%) reported the

results of a CUA and a CBA, and eight studies (32%) conducted a CEA, CUA and CBA. Five methods were used to estimate utilities, with the EuroQol five dimensions questionnaire (EQ-5D) being the most popular. A wide variety of methods were used to monetise benefits within a CBA, including the human capital and friction cost approaches, DCEs, WTP questionnaires, the value of a statistical life, and the contingent valuation approach.

Only five studies (20%) provided explicit justification for adopting a welfarist approach. Of these, three were concerned that the extra-welfarist approach would not capture important elements of benefit. One noted that a CUA approach may underestimate the benefits of cycling interventions by ignoring externalities [70]. A second commented that QALYs may not capture all of the benefits of varicella vaccination, as compared to a WTP approach [71]. A third anticipated that WTP would be a more sensitive method for evaluating non-life-threatening disorders [72]. The fourth study argued that CBAs are the best way to decide whether a healthcare intervention is worth implementing [73]. The final study used EQ-5D valuations based on population data in a CUA but undertook an additional CBA due to concerns that these valuations may not match the preferences of patients [25]. Justification for adopting a welfarist approach was not found to be related to the intervention context (i.e. whether or not it was a public health intervention).

Seven studies (28%) reported results that showed limited or no concordance. Only one of these studies was commercially funded (a comparison of different bronchitis treatments [74]) and only one evaluated a public health intervention (varicella vaccination [71]). Four of the seven studies (57%) were undertaken in the UK [25,26,71,73]. There were few clear links between the methods used to estimate utilities or monetise benefits and the

concordance of the results. One exception was that all studies using the human capital approach reported full concordance in their results [40,75-80]. In addition, both studies which conducted a CBA using DCE data reported no concordance in their results [25,26].

In four of the studies showing limited or no concordance, the authors judged that extra-welfarist approaches provided the most convincing evidence for adoption. Two studies in psoriasis treatment and back pain management concluded that all treatments were cost-effective, but not cost-beneficial [72,81]. A third study evaluated ciprofloxacin versus usual care for bronchitis. According to the authors, the CEA and CUA results provided evidence of value for money, but the CBA results were not favourable [74]. A fourth study of hernia surgery reported that one specific laparoscopic approach was cost-effective in a CUA, but that no laparoscopic surgery approaches were associated with net benefits in a CBA [25]. In two further studies showing limited or no concordance, it was unclear which approach provided the most convincing evidence for adoption. A study evaluating deep brain stimulation versus no intervention in neurosurgery concluded that neither a CUA nor a CBA provided strong evidence for adoption [73], while a second study evaluating surveillance strategies for ocular hypertension found that the most efficient strategy was different in all three forms of economic evaluation (CEA, CUA, CBA) [26]. In the final study showing limited or no concordance the welfarist approach was judged to provide the most convincing evidence: the economic desirability of a varicella vaccine was uncertain using CUA but beneficial using CBA [71].

Only three studies (12%) identified the most informative economic evaluation results. In an economic evaluation of diet therapy versus drug therapy for the treatment of

hypertension, the authors noted that CBA is theoretically better able to handle such decision problems, as patients consider the costs of side-effects and leisure time when stating their WTP [82]. A second study focused on the CEA and CUA results, expressing concerns over the methods used to assign a monetary value to life and quality of life [83]. In both studies, there was concordance between the economic evaluation results. A final study (in which there was limited concordance) also favoured the CEA and CUA results (which support the use of ciprofloxacin to treat bronchitis), with the authors declaring that they had little faith in the CBA results (which favoured usual care) [74].

The assessment of study quality (**Appendix 2**) classified seven studies (28%) as low quality, 13 (52%) as average quality, and five (20%) as high quality. Overall, studies met 63% of quality criteria, suggesting the overall body of evidence was of average quality. Criteria which were met by few studies included providing justification for the choice of decision-model and appropriately characterising heterogeneity. There was some evidence of an improvement in quality over time. In addition, studies indicating limited or no concordance scored more highly (71%) than those indicating concordance (60%).

Discussion

The results of the literature review indicate that for every four studies applying both the welfarist and extra-welfarist approaches, one shows limited or no concordance between the results of the different forms of economic evaluation, suggesting that choice of approach can impact on adoption recommendations. The seven studies reporting limited or no concordance evaluate interventions across a variety of country settings and disease areas (although studies evaluating public health interventions are more likely to show

concordance), and the economic evaluation approach providing supportive evidence for adoption varies between these studies. There is some limited evidence from four of these seven studies which suggests that when full concordance is not achieved, it is the extra-welfarist results that may provide the most convincing evidence for adoption.

When results from multiple approaches are reported simultaneously, researchers provide insufficient guidance to enable decision-makers to make fully informed adoption decisions. Only three studies identify the most informative economic evaluation results, and only one of these is a study in which there was limited or no concordance. Furthermore, few authors explain why both welfarist and extra-welfarist economic evaluations have been conducted. Given that decision-makers favour the extra-welfarist approach, and there are fundamental theoretical differences between the two approaches, this is surprising.

The source of funding was identified for each study, on the basis that commercially funded studies might be more likely to apply a variety of economic evaluation approaches (in order to more effectively state a case for adoption), and might be more likely to return a finding of concordance. However, no evidence was found to support either hypothesis. There was, however, evidence of a link between a finding of no concordance and higher study quality. The overall body of evidence was of average quality. One potential explanation for this finding could be that journal word limits preclude the full reporting of the methods of studies which conduct more than one form of economic evaluation.

Fewer references than anticipated were identified that reported the results of both welfarist and extra-welfarist economic evaluations. In part, this is because cost

comparisons are often mislabelled as CBAs, creating the impression that more CBAs have been conducted than is actually the case. This result has been noted previously [58,84-86], and is compounded by the fact that MeSH terms other than 'cost-benefit analysis' have only been made available relatively recently [87]. A second explanation is that welfarist and extra-welfarist economic evaluations conducted using the same dataset or model may be reported in separate journal articles, and both references will not necessarily be captured by the search strategies used within this review. This is a limitation of this review, although given funding and time constraints, few health economists are likely to be able to conduct more than one form of economic evaluation within the same study.

Ten of the references included in this review were identified from other sources, which could be an indication that the search strategies were not specified sufficiently accurately. This issue was anticipated prior to undertaking the review, given the use of quite general search terms. An exploration of alternative literature sources was therefore planned alongside the main review. When these additional references were reviewed no alternative search terms were identified that could have improved the main review. Other strategies were considered, such as searching only for CBAs and then performing focused searches for corresponding extra-welfarist economic evaluations. However, given the widespread misclassification of CBAs within the health economics literature, it was thought unlikely that such a strategy would identify additional studies reporting both welfarist and extra-welfarist economic evaluations.

Overall, the finding that one in four comparative studies shows limited or no concordance appears to be reasonably robust, bearing in mind the above caveats. A key consequence of

this finding is that incorrect adoption decisions might be made if an inappropriate economic evaluation approach is used. This may be a particular problem for complex interventions such as genomic technologies. However, this is not the only methodological challenge that is faced when conducting economic evaluations of such technologies. The following section takes a closer look at these additional challenges.

2.4 Methodological challenges in genomics

In Chapter One it was noted that translational research evidence for genomic technologies is rare, particularly economic evidence. One reason for this lack of evidence is the uncertainty surrounding the appropriate economic evaluation approach in this context, and the extent to which choice of approach can impact on adoption decisions was discussed in the previous section. However, additional methodological challenges also exist when conducting economic evaluations of genomic technologies, and this thesis aims to address several of these challenges. This section presents a literature review which summarises these challenges.

2.4.1 Literature review aims, objectives and methods

The aim of the review was to identify references which considered the methodological challenges faced when conducting economic evaluations of genomic interventions. As the terms 'genetic' and 'genomic' are often used interchangeably in the literature, both were included in the literature searches [88,89]. References were included in three categories:

1. Methodological health economics references considering the challenges associated with conducting economic evaluations of genetic or genomic technologies;

2. Applied health economics references highlighting methodological issues when reporting the results of economic evaluations of genetic or genomic technologies;
3. References considering the translation of genetic or genomic technologies into clinical practice and describing methodological issues that could arise from a health economics perspective.

It was anticipated that a number of non-health economics references in category three would meet the inclusion criteria. The review was therefore designed to capture these references by combining multiple keyword variations in OR logical statements, maximising sensitivity and minimising false negatives. Literature searches were undertaken in Embase, Medline, Econlit, NHS EED, the UK NHS HTA Programme, and The Cochrane Library. Keywords were identified by examining relevant references in the literature, previous related literature reviews, and MeSH terms. Searches were limited to full length references published in English (to enable in depth reviews to be carried out), references published from 1990 onwards (to ensure that results were applicable to the present setting), and studies in humans. Details of the search strategies are provided in **Appendix 5**.

Searches were undertaken in April 2013. References which met the inclusion criteria were retrieved and information was extracted on methodological challenges falling into one of five categories: Analytical Approach, Costs and Resource Use, Measuring Outcomes, Measuring Effectiveness, and Other.

2.4.2 Literature review results

Figure 2.3 summarises the literature search results. 2772 references were identified, with 1868 titles remaining after removing duplicates. Of these, 30 met the inclusion criteria. References which evaluated pre-implantation and prenatal genetic and genomic interventions for family planning purposes were not included: these interventions present several specific challenges for health economists (e.g. should the projected benefits from the termination of affected foetuses be included in economic evaluations?), which have been discussed elsewhere [90,91]. 22 references were added from other sources (e.g. earlier literature reviews), so 52 references in total were included in the review (**Appendix 6** presents a complete list). Of these, 39 were published from 2008 onwards.

The methodological challenges that were identified in the literature search are described in the following sub-sections and summarised in **Table 2.2**, which also describes potential solutions. Since the review was undertaken a number of additional relevant articles have been published by key authors in this field. These articles are not included in the paper counts for the review, but are referenced in the remainder of this section as appropriate.

Figure 2.3: Literature review results – methodological challenges faced when conducting economic evaluations of genomic interventions

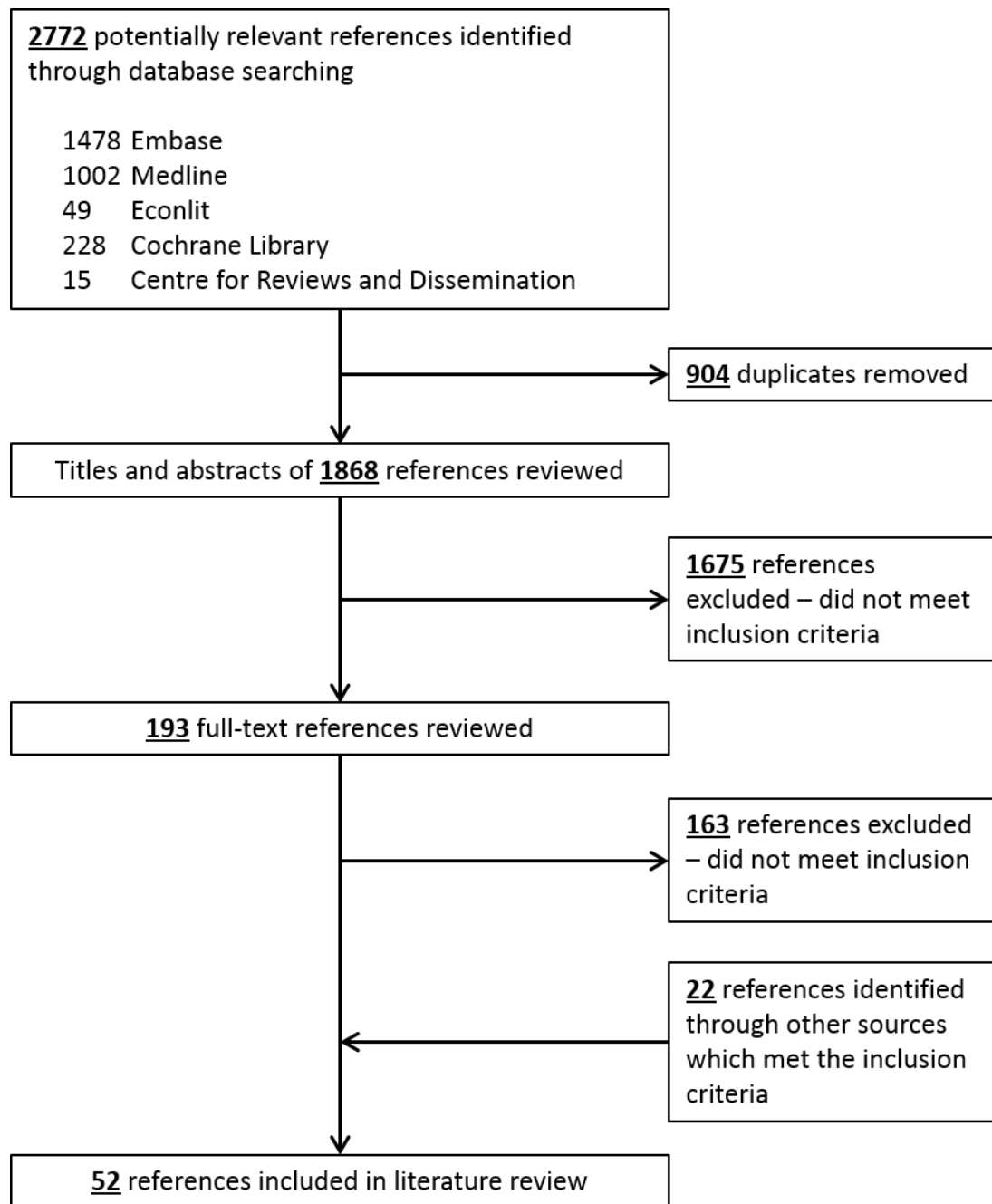


Table 2.2: Summary of methodological issues arising in health economic evaluations of genomic technologies

Category		Summary of issues	Possible solutions
Analytical Approach	1	The correct perspective to use remains uncertain.	Future studies should consider multiple analytical perspectives. As costs and outcomes are likely to be accrued over long time horizons, discount rates should be carefully selected. Analyses should be undertaken in an iterative manner and rerun when new evidence becomes available: this may require extensions in health economic modelling capacity. Study designs must take into account all potential comparators (both genomic and non-genomic) and all subsequent therapeutic decisions. In general, realistic simulations and comprehensive parameter and structural sensitivity analyses should be standard practice.
	2	The appropriate analytical timeframe can be many years in length. Studies that focus on the short-term costs and consequences of genomic interventions risk misestimating cumulative costs.	
	3	The timing of an economic evaluation of a genomic test is critical as standard testing practice evolves continuously.	
	4	The choice of comparator and the specification of the study design can significantly impact on economic evaluation results.	
Costs and Resource Use	1	A large number of cost categories are potentially relevant when conducting an economic evaluation in this context, including the costs of patient recruitment, sample collection and testing, data analysis, communication of test results, and actions taken based on test results.	A much broader range of costs must be included in genomic economic evaluations, with these costs collected at more frequent intervals. Analyses must also be increasingly flexible in order to account for temporal and geographical price variations. Costing studies of platform diagnostics with multiple applications should be conducted particularly carefully.
	2	There are no national pricing tariffs for genomic tests. Costs vary considerably both within and between countries.	
	3	The combination of unstable tumour genomes and evolving test filters may require genomic tests to be conducted increasingly frequently.	

Category		Summary of issues	Possible solutions
Measuring Outcomes	1	Disease specific and preference based outcome measures limit comparability and do not capture all relevant dimensions of outcomes.	Studies that evaluate outcome measures in this context which are not disease specific or preference based would be useful. In particular, methods that can be used to value outcomes from genomic interventions within a cost benefit analysis framework would be particularly valued. As individual outcomes may be more important than population outcomes, subgroup analysis is likely to be crucial.
	2	Capturing information on personal utility may be important, but the metrics for measuring personal utility are not well established.	
	3	Cost benefit analyses might have a greater role to play in this context than other forms of economic evaluation.	
	4	Individual outcomes may be more important than population outcomes.	
Measuring Effectiveness	1	Little is known about how patients and clinicians behave when faced with information provided by genomic interventions.	Behavioural uncertainty should be incorporated into analyses. The individualistic nature of treatment should also feed through to study design. There is a greater need for post-implementation economic analyses and evaluations, based on pragmatic trials. Finally, studies that consider the use of summary measures such as polygenic risk scores within economic evaluations are likely to be valued, as well as studies that better link genomic data with health outcome measures.
	2	The quality of effectiveness data is weak.	
	3	Effectiveness data for genomic interventions, when available, is complex and challenging to incorporate into standard health economic analyses.	
Other	1	Performance standards for genomic interventions vary considerably between laboratories.	Analyses should consider all possible models of service delivery.
	2	The pace of innovation in genomics suggests that prioritising investment in expensive yet informative comparative studies is desirable.	Value of information analysis is an important component of economic evaluations of genomic technologies.

Analytical approach

Perspective

Most economic evaluations of non-genomic interventions consider the impact of an intervention on the health service within which it is used, as recommended by HTA agencies. In contrast, a societal perspective considers indirect impacts such as the financial costs incurred by families due to an intervention. Some HTA agencies (e.g. Italy, Norway) permit the submission of additional analyses to accommodate a wider societal perspective, provided that these additional results are presented in a disaggregated manner [92]. This wider perspective may be beneficial in genomics as testing can affect both healthcare choices and broader life decisions (e.g. choosing a child's school). However, at present, only 27% of genomic economic evaluations adopt this perspective [93,94].

Timeframe of analysis

The information provided by genomic interventions can have long-term implications for patients (e.g. identifying the presence of mutations linked to Lynch syndrome in asymptomatic patients [95]). Consequently, studies focusing on short-term costs and consequences risk underestimating the cumulative costs and effects of these genomic interventions [18,91,93,96-98]. The time horizon of a genomic economic evaluation should therefore be carefully selected to reflect all of the costs and benefits which may occur in a patient's lifetime, not just those incurred at the point of testing.

Timing of analysis

Genomic testing practice changes continually, and the costs, effectiveness and outcomes of new genomic technologies are particularly poorly defined [66,88,93,96,99-101]. For

example, the cost of sequencing has fallen dramatically in just a decade [102]. Patient categorisation can also vary over time as additional subgroups are identified with implications for individualised treatment (e.g. type 2 diabetes risk assessment using genotyping) [103-106]. Furthermore, many genomic tests (e.g. the use of microarrays for the diagnosis of idiopathic learning disability) gradually emerge into clinical practice from a research setting, instead of entering the healthcare system formally via regulators [18]. This makes it difficult to compare economic evaluations conducted at different points in time [93,99,101].

Uncertainty concerning the appropriate timing of an economic evaluation does not necessarily mean that an evaluation should not be undertaken. Indeed, economic evidence is valued by decision-makers at all stages in the translational process, particularly the early stages of product development when decisions about prioritisation of research funding are being made, and there is likely to be a role for techniques such as the Headroom Method (a form of early economic evaluation which considers the likely commercial viability of an intervention) when deciding whether to proceed with the development of a genomic test [96,107,108]. Modelling the potential impact of a learning curve on both unit costs and effects could also be useful, provided that these estimates are then validated by collecting cost and resource use data at the end of a study period to allow for any reductions in price or in rates of testing errors.

Analytical content

Genomic tests often have multiple applications in varied contexts and it is rarely possible to synthesise results from economic evaluations undertaken in these differing settings

[98,109]. Furthermore, different applications may not be similarly cost-effective (e.g. Oncotype DX cost-effectiveness differs between breast and colon cancer [110,111]). This means that study design must be carefully considered, with particular attention paid to two elements. First, standard genomic testing practice rarely exists; it varies considerably both between and within clinical specialties [112]. It is therefore important to carefully choose the comparators in a genomic economic evaluation. It may be appropriate to evaluate genomic testing both against, and in addition to, non-genetic testing [113,114]: For example, expensive yet specific genomic tests are more cost-effective in colorectal cancer if combined with preliminary screening using cheaper less specific tests [101,109,115,116]. Second, study designs should take into account the therapeutic decisions that follow from the use of a genomic intervention [99]. For example, genetic testing for Duchenne muscular dystrophy can detect mutations across 79 genetic regions, with 30 different therapies required, each targeted at specific faulty regions. In this scenario it is insufficient to conduct an economic evaluation of just one test specification: costs and outcomes will vary depending on mutation prevalence and treatment efficacy [117]. Multiple potential test designs exist, and analysts must carefully consider which comparators are appropriate prior to undertaking an economic evaluation.

Costs and resource use

Which costs should be included?

The wide range of potential costs in genomic economic evaluations requires health economists to carefully plan data collection to ensure that all relevant costs are included. Such costs may include those associated with patient recruitment (e.g. publicity and patient education) [91,93], sample collection, laboratory testing [93], data analysis (e.g.

bioinformatics) [88,118], and reporting test results to patients [93]. The costs associated with actions taken on the basis of test results are also important, including the costs of treatment and follow-up testing (or tests avoided), management of adverse events, and genetic counselling [91,93,119-122]. These costs can be significant. A study comparing microarray testing with karyotyping in idiopathic learning disability found that including all the costs of follow-up testing after karyotyping changed the study conclusions in favour of microarray testing (the more expensive test on a standalone basis) [123]. The time and cost burden of monitoring disease progression and drug response may also be important: the most cost-effective genomic tests are likely to be those for conditions where monitoring is difficult and expensive [122].

Depending on the analytical perspective chosen, the indirect costs accruing to patients may also require estimation (e.g. time spent seeking treatment, time lost from work due to poor health). For example, if genomic tests can identify patients who will not benefit from treatment, these patients may be spared the costs associated with inappropriate treatment, including work absenteeism. However, few studies have considered these costs: a study evaluating the cost-effectiveness of Oncotype DX in breast cancer is one exception [124].

Finally, it may be appropriate to account for the wider infrastructure costs associated with genomic technologies (e.g. staff training costs). However this can be difficult to achieve in practice as genomic platform diagnostics often have multiple applications: a piece of equipment may initially provide one specific test, but ultimately develops into a 'one stop shop' for genomic testing in a variety of clinical areas [66,88]. In this scenario, the

appropriate allocation of infrastructure costs is not clear: initial applications of a genomic technology may be burdened by a disproportionate share of overall infrastructure costs, while later applications benefit from economies of scale realised through more intensive use. This may bias economic evaluation results.

How much do genomic interventions cost?

National guidelines and agreed reimbursement rates (e.g. NHS tariffs, Medicare payment rates) are generally widely available for commonly used treatments and procedures in countries with HTA agencies. However, these costs are rarely available for genomic tests. Some cost estimates are available from individual testing laboratories, but these can vary widely and are not always accompanied by resource use data [18,19,88,121,125,126]. Cost differences between tests developed by accredited laboratories and commercially marketed tests can be particularly acute [119]. Costs also vary between countries due to local price differences, variations in practice, and differences in patient contributions to test costs [18,127]. This makes it difficult to conduct economic evaluations of genomic interventions as many HTA agencies (e.g. NICE) strongly recommend that nationally agreed prices are used within their technology appraisal process. More generally, tests are rarely assigned prices that are value-related [3]. This issue will become increasingly problematic over time as more tests consider multiple genetic changes in a single assay.

When should cost data be collected?

Frequency of data collection depends on the type of genomic intervention being evaluated. Tests that measure germline genetic changes may only need to be performed once [128]. In contrast, multiple genomic tests may be required to identify changes in tumour genomes

which are acquired over time (e.g. in chronic lymphocytic leukaemia (CLL) [129]). In addition, genomic tests are often packaged with bioinformatics packages which are regularly updated to incorporate new research, hence results may need to be retrieved and re-analysed multiple times, incurring additional costs [91,130]. Furthermore, as whole genome sequencing becomes more widespread, it may become cheaper and easier in the long-term to sequence a patient's entire genome multiple times, rather than retrieving and re-analysing data for new indications, presenting an additional cost complication [88]. In contrast, many commonly used non-genomic tests (e.g. an x-ray of a suspected broken bone) only need to be performed once to provide definitive answers, simplifying decisions regarding the collection of cost data.

Measuring outcomes

Disease specific and preference-based outcome measures

Disease specific outcome measures are occasionally informative in genomics [119], but in general they do not capture all relevant dimensions of outcomes, in particular non-health outcomes (e.g. process-related outcomes such as improvements in waiting time). This is a particular problem for genomic interventions which provide only diagnostic information, which can reduce anxiety and help patients to make future plans. Generic preference-based measures such as the EQ-5D improve comparability by collecting data across a broad range of health related quality of life domains. However, QALYs measured using such metrics do not capture family spillover effects [15] or non-health outcomes (e.g. the value of possessing genomic information) [15,88,91,93,97,101,106,119,120,131-134], and thus cannot reflect all possible health states for genomic interventions (e.g. living with the knowledge that you have a mutation) [115]. As some genomic interventions do not

improve health or extend life (e.g. microarray testing in learning disability provides a diagnosis but does not lead to a treatment), preference-based measures may not pick up differences in outcomes, resulting in high ICERs which suggest that these tests are poor value [62].

Personal utility

Genomic interventions do not always provide clinical utility, but may impact on patient well-being via non-clinical routes [18,62,65,91,93,106,118,135]. Such effects are captured by the term 'personal utility' (*'benefits or harms that are manifested primarily outside medical contexts'*) [133,135]. Factors that can have a positive effect on personal utility include the acquisition of prognostic or diagnostic information that can enhance individuals' sense of control, thus improving personal accountability for health-related choices, informing a sense of self-identity and autonomy, and potentially improving uptake and adherence [66,91,135,136]. Genomic information may also provide reassurance and relieve anxiety [96,109,137,138], allowing patients and relatives to make lifestyle modifications (e.g. APOE testing in Alzheimer's Disease) [62,96,109,139], and reducing the frequency of diagnostic odysseys (e.g. idiopathic learning disability [123]). Finally, genomic interventions can provide healthcare process related benefits such as improvements in waiting time, as test results can be returned more quickly [65].

Genomic interventions can also have a negative impact on personal utility. Anxiety may be increased if a test suggests treatment non-response (leaving a patient without the prospect of disease amelioration) [101,140], or if secondary findings with clinical implications are reported [62,91,132,135]. Genomic test results can also disrupt family dynamics and

complicate reproductive decisions [91,132,135], lead to fear of discrimination and stigmatisation (if an adverse result is discovered) [15,98], or, conversely, lead to complacency, encouraging unhealthy behaviours if a test result appears favourable [15].

Unfortunately, the metrics for measuring personal utility are not well established [99,109,139]. Work is underway in the UK to consider how the current definition of value used by NICE could be expanded to consider personal utility and other factors [66], but the results of this work have not yet been published. Some studies have developed outcome measures based on particular concepts of personal utility (e.g. empowerment [62]), but these approaches have limitations [103]. The capability approach – which suggests that wellbeing should be measured by what people have the potential to do (their capabilities), not what they actually do (their functioning in daily life) – may also be useful for capturing information on personal utility, but this has not been applied widely in genomics [62].

Cost-benefit analysis

The key methodological disagreement amongst health economists is how to measure the health outcomes associated with genomic technologies, with many now advocating the more widespread use of a welfarist approach in order to incorporate personal utility into economic evaluations [15,93]. The wider context surrounding this debate has been described in detail earlier in this chapter. Several of the approaches used to value outcomes in monetary terms in CBAs may be well-equipped to capture the non-health outcomes of genomic testing [15,62,66,88,91,93,98,106,119,132]. Stated preference approaches such as DCEs may be particularly appropriate, but DCEs have rarely been used in genomics, or to inform CBAs, possibly because people struggle to make hypothetical

choices in unfamiliar contexts [15,93,139]. An alternative approach which may avoid some of the problems associated with DCEs is best-worst scaling (BWS) [141]. However, only one study, which evaluated physician preferences for personalised medicine, has applied BWS in a genomic context [142]. This lack of evidence from DCEs or BWS studies is a clear obstacle to the more widespread application of the welfarist approach in this context, but this situation may not persist: the US Food and Drug Administration has recently issued guidance which recommends that patient preferences are incorporated into the decision-making process for medical devices in the USA [143].

Given these drawbacks, and the fact that CBAs are rarely accepted by HTA agencies, it has been suggested that analysts may also wish to consider alternative approaches such as cost-consequence or comparative effectiveness analysis [88], or the use of the capability approach within an extra-welfarist framework. However, these approaches have also been applied infrequently to date and may not be accepted by HTA agencies either.

More generally, health economists should carefully consider how the genomic intervention that is being evaluated operationalises the personalisation of medical treatment as this can influence the choice of evaluative framework. Rogowski *et al.* suggest that if physiological biomarkers are being used to stratify healthcare interventions to patient subpopulations, conventional (extra-welfarist) evaluative frameworks will generally be appropriate. However, if stratification is informed by individuals' preferences for elements of health and related wellbeing, evaluative frameworks may need to be adapted [144].

Individual versus population outcomes

It is important to consider patient-level outcomes when evaluating genomic tests. Where such tests inform treatment decisions, response to treatment will vary considerably by patient [101,135,137,145], and patient subgroups will have different opinions on acceptable levels of test sensitivity and specificity [146]. These individual differences in risk perceptions and attitudes are to be expected, given the complexity of genomic test information [101]: one study found that the value placed on genetic testing for colorectal cancer varied significantly according to patient gender, income and education [147]. Consequently, adopting a narrow 'population' approach will likely obscure differences in cost-effectiveness across patient subgroups. However, converting the evidence base from individual tests into an understanding of the value of obtaining the same information in aggregate is difficult [118]. Pre-screening activities (e.g. taking a family history, counselling to identify the likely impact of genomic information on patients) may improve the cost-effectiveness of these interventions and such activities could be incorporated into economic evaluations, if this is judged to be appropriate [113,135].

Measuring effectiveness

Patient and clinician behaviour

The cost-effectiveness of genomic interventions will depend on several factors including whether patients and clinicians use these interventions and how they respond to test results (e.g. whether patients undergo appropriate treatment). Ideally, these factors should feed into economic evaluations [91,96,101,137,138], but little data on behaviour is available [101]. Those studies that do exist are unclear on whether all patients will want to be tested [148], whether they will adjust their behaviour to fully adhere to test advice

[97,98,125,126], and whether physicians will take genomic information into account when making clinical decisions [101]. Limited evidence suggests that if testing is not mandated, universal uptake is unlikely [121].

Effectiveness data quality

The quality of effectiveness data for genomic tests is often weak [3,18,99,106]. Many genetic conditions are uncommon, and large randomised controlled trials (RCTs) would be required, running over many years, at great expense, to generate sufficient power to detect clinically significant endpoints. Given the rapid evolution of genomic technologies, RCTs may not even provide relevant results [66,88,91,97,101,105,109,130,145,147,149]. Compounding these issues, tests used in routine clinical practice often do not perform as well as in research settings [114], and test performance can vary depending on laboratory characteristics (e.g. size [121], choice of equipment [100]).

Consequently there have been calls for a more dynamic trial environment in genomics, utilising more flexible clinical trial designs [150]. Non-inferiority trials may be an alternative source of effectiveness data. However genomic interventions may supplement rather than replace existing practice, and non-inferiority trials may need to be larger than RCTs designed to show superiority [140]. The personalised nature of genomic interventions means that to capture their true effects, studies must focus on patients who are outliers [128]. A more individualistic approach to the measurement of effectiveness may therefore be required (similar to N-of-1 clinical trials) [7,88], although it is not easy to conduct prospective well-controlled studies of these outlier patients as they are so rare [128].

Regulators are increasingly open to the use of effectiveness data from alternative sources in this context, including disease registries, observational and cohort studies, data generated in community settings, and expert opinion studies [3,65,66,101,120,125]. However, these data may suffer from the problem of inadequate numbers [109,114], such studies cannot account for recent advances in treatment [97], and expert opinion estimates may be biased [120]. Post-implementation analyses could instead provide the robust effectiveness evidence that is required [66,125,138,140,149,151]. However, such evidence may only have a limited impact on coverage decisions, as regulators rarely sanction disinvestment in technologies that are no longer considered to be sufficiently effective. Bayesian (or adaptive) trial designs may be able to evaluate fast evolving genomic technologies, but regulators are not yet convinced by these new methods [7,99]. For genomic interventions already in clinical practice, large administrative databases could be used to access routinely collected clinical data and to link this to genomic data (e.g. Biobank data). Such databases provide information on both infrequent events and routine care [66,99,149,151], but detailed information on specific genomic interventions may be limited.

It is becoming increasingly apparent that the methods that are currently used to generate effectiveness evidence for economic analyses are not readily scalable to the pace of genomic discovery in terms of turnaround time and expense [66,118]. ‘Wait and see’ policy conclusions may become increasingly common in economic evaluations of genomic technologies as decisions are delayed until better evidence emerges [88].

Data complexity

The information provided by genomic technologies is challenging to incorporate into economic evaluations as the evidence base linking complex genomic data to health outcome measures is generally underdeveloped [19,118,134]. For example, there are over 130 variations of the *CYP2D6* gene (which indicates tamoxifen resistance), but it is unclear which are the most clinically relevant [128]. More generally, genomic testing outcomes are often influenced by multiple genes, each genetic variant can influence multiple outcomes, and the influence of a variant on a given outcome can vary across individuals [101]. This makes it challenging to structure a model-based economic evaluation which evaluates a test for multiple biomarkers across several contexts, and may necessitate a move away from commonly used Markov cohort model approaches to microsimulation or discrete-event simulation approaches [144].

Further methodological issues include ‘multiple testing’ (how to control the false positive rate) and the ‘winner’s curse’ (the effect of a newly identified mutation is commonly overestimated in initial studies) [99]. It is also currently unclear how to incorporate information on diagnostic characteristics (e.g. receiver operating curves) into economic analyses. Compounding these challenges is the fact that the performance of commercially developed and ‘in house’ genomic interventions may differ considerably [65,120].

One way in which clinical utility can be extrapolated from complex genomic data is to apply multiple cut-off points to identify clinically relevant subgroups with similar genomic profiles [120]. Alternatively, polygenic risk scores can combine all available genotypic data in one measure, attaching a score based on the number of mutations (e.g. genomic

complexity in CLL [129]). However, such scores may require large sample sizes to reach statistical significance [152]. A further issue is that separate studies will rarely converge on identical scoring tools, making it difficult for decision-makers to synthesise evidence from multiple studies [93].

Other

Service delivery

Performance standards for genomic interventions vary considerably between laboratories, hence choice of service delivery model should be carefully considered by health economists. For some disorders (e.g. HER2 testing in breast cancer) laboratories with lower testing volumes are more likely to report incorrect findings, suggesting that models of service delivery focused around specialist regional testing hubs might yield more favourable economic evaluation results [66,99,148,149].

Research prioritisation

Value of information (VOI) analysis is increasingly used to calculate the expected costs and benefits of future research in healthcare, and may be an important component of economic evaluations of genomic technologies [98,99,101,120,133,153]. VOI techniques allow researchers to highlight areas where new evidence is likely to be valued by decision-makers, and the pace of innovation in genomics suggests that prioritising investment in expensive yet informative comparative studies is desirable [99]. Furthermore, if evidence of the effectiveness of a genomic intervention is weak, the costs and consequences of collecting additional data to better inform translational research must be carefully considered, particularly if RCTs are not feasible [97,154].

Only one study has applied these techniques in a genomics context to date [153] and there are several issues to be addressed before these approaches can be more extensively applied in the future. A key problem is that the populations that may benefit from these interventions are small, so low cost approaches to evidence generation are required to ensure that the benefits derived from new evidence surpass the research costs. VOI analysis also incorporates an assumed time horizon for the research prioritisation decision which is challenging to estimate in genomics. Although the pace of technological progress is rapid, inferring a relatively short product lifecycle, this contrasts with the fact that when new biomarkers are discovered they are often complements to existing knowledge rather than substitutes, and full validation of these biomarkers will take time, inferring a longer time horizon. These two effects add considerable complexity when selecting an appropriate time horizon [144].

2.4.3 Discussion

This literature review confirms that economic evaluations of genomic technologies pose several methodological challenges for health economists. The appropriate analytical perspective, timeframe and timing are all uncertain, the range of costs that could be included is wide, and tests may need to be repeated an uncertain number of times over an undefined time period. Commonly used outcome measures may not be appropriate as they may not capture all relevant dimensions of outcomes. However, measures which focus on personal utility (which are potentially more informative), are not yet well-established. Welfarist approaches to economic evaluation may have a greater role to play, but instruments such as DCEs have rarely been used in genomics or to inform CBAs. Furthermore, little is known about how patients and clinicians behave when faced with

information provided by genomic interventions, and this is rarely factored into economic evaluations of these technologies. Effectiveness data are commonly weaker than the data which exists for non-genetic interventions, and large RCTs are rarely feasible. Finally, it can be tricky to incorporate complex genomic effectiveness data into economic analyses: the evidence base linking this data to health outcomes is underdeveloped, and possible solutions (e.g. polygenic risk scoring) are still under development.

Some of these challenges are specific to genomics, including those associated with variations in genomic test performance standards between laboratories, identifying an appropriate allocation of infrastructure costs for genomic platform diagnostics with multiple applications, and the difficulties surrounding the selection of an appropriate time horizon in VOI analysis. However, many of these methodological issues also arise in economic evaluations of non-genomic technologies. Indeed, a number of these challenges (e.g. learning curve effects, unstable prices, incremental innovation) have been highlighted in the literature on the economic evaluation of medical devices [155]. Consequently, although it is clear that economic evaluations of genomic technologies present a particular challenge for health economists, it is the accumulation of methodological challenges in this context that presents the greatest difficulty, rather than the precise nature of these challenges.

It could be argued that resolving some of these issues is the responsibility of researchers working in other disciplines (e.g. evaluating the clinical utility of genomic technologies in the face of their rapid evolution). However, health economists must still be prepared to adapt their methods to account for these challenges, and studies published since this

review was undertaken have come to similar conclusions [156,157]. A number of potential solutions have already been proposed which could be incorporated in future genomic economic evaluations (summarised in **Table 2.2**). Some are straightforward (and not necessarily specific to genomics): realistic simulations and comprehensive sensitivity analyses should be standard practice to allow health economists to fully consider the impact of different analytical perspectives and test characteristics [18,93]. Other solutions require an evolution in methods. Health economists will need to work in a more iterative manner, quickly incorporating new evidence into analyses and re-running simulations regularly [66,88,93,101,103,105,106]. Early-stage prospective economic evaluations will become more common [120,121,134] and genomic effectiveness data will increasingly originate from studies which are not RCTs. ‘Wait and see’ or ‘coverage with evidence development’ conclusions are likely to become more frequent as a result. VOI techniques will have an important role to play, as generating the evidence required to inform resource allocation decisions in genomics will require careful prioritisation of limited healthcare research funds.

The area where there is most disagreement is how to measure the health outcomes associated with genomic technologies. Some maintain that *“the ‘new genetics’ does not pose new problems for health economics, but it highlights aspects of evaluation that have been neglected in previous economic evaluation research”* [132]. Others claim that ‘genetic exceptionalism’ exists, and new methods are required to evaluate the outcomes arising from genomic technologies [91]. Several authors now advocate a welfarist approach in order to incorporate information on personal utility into economic evaluations [15,93]. However, it was noted earlier in this chapter that health economists are somewhat

constrained in their approach to outcome measurement by the regulatory context within which economic evaluations of genomic technologies are conducted. Ultimately, incorporating personal utility into economic evaluations will depend on the scope of decision-makers. For regulators deciding which tests should be available, it may be reasonable to consider personal utility alongside other dimensions of utility. From a public healthcare provider perspective (with an overall aim of improving health), personal utility may not be relevant [15,96].

2.5 Conclusions

This chapter began by providing an overview of the rationale for economic evaluation before considering how the debate concerning the most appropriate approach has evolved over time. Two literature reviews were then presented that placed the research questions which will be addressed in this thesis into context. The key result from the first of these reviews was that for every four studies applying both the welfarist and extra-welfarist approaches, one shows limited or no concordance between the results of the different forms of economic evaluation. This suggests that choice of approach can impact on adoption recommendations (whether this affects actual adoption decisions will depend on the extent to which healthcare decision-makers consider economic evaluation results in their decision-making process). This may be a particular problem for complex interventions such as genomic technologies. The second review then confirmed that the quantity and range of methodological challenges is such that economic evaluations of genomic technologies present a particular challenge for health economists.

Taken together, these literature reviews provide the methodological background to this thesis, and these findings will be taken forward in the following chapters. To address the primary research question of this thesis (“What is the most appropriate economic evaluation approach to use in genomics?”), both the welfarist and extra-welfarist approaches will be applied to tackle the same decision problem, namely what are the costs and consequences associated with the use of genomic testing to guide chemotherapy treatment decisions in CLL patients in the UK? These different approaches will also consider some of the methodological challenges highlighted in the second literature review. The consequences of applying different analytical perspectives and timeframes will be considered, and data will be presented on the cost of genomic testing in CLL. The two approaches will also incorporate both complex genomic effectiveness data and information on personal utility, and VOI analysis will be used to quantify the costs and benefits associated with further research in this context. These analyses will be presented in Chapters Four, Five and Six. For now, the attention of this thesis will turn towards Chapter Three, which describes the case study disease area and the genomic technology which will be evaluated in these analyses.

Chapter Three | Chronic lymphocytic leukaemia and genetic testing

3.1 Introduction

This chapter has three aims. First, it will provide an overview of current patient management and treatment practice in the case study disease area considered in this thesis: chronic lymphocytic leukaemia (CLL). Second, it will describe the rationale for using genomic testing to guide treatment decisions in CLL. Third, the case study technology (targeted next generation sequencing) evaluated in this thesis will be described, and the economic evidence for the use of this technology in all disease areas will be summarised.

3.2 Chronic lymphocytic leukaemia

This section describes the epidemiological basis of CLL and provides information on typical practice regarding presentation, diagnosis and staging. Given that information provided by genetic testing can inform treatment practice, both current and likely future testing practice are summarised, and the health economic evidence on genetic testing in CLL is reviewed. Finally, the clinical management of CLL patients in the UK after the point of diagnosis is described.

3.2.1 Epidemiology

Approximately 3,500 people are diagnosed with CLL in the UK annually [22]. CLL is the most common adult leukaemia, and accounts for 25-46% of all leukaemias [20-23]. The UK incidence rate is four cases per 100,000 per year, increasing to 30 cases per 100,000 in those aged 80 and over [158,159]. Variations in incidence by ethnicity have been observed:

higher rates are reported in Caucasians than in people of Asian or South American ethnicity [158,160-163]. In addition, two-thirds of CLL cases are in males [21,22,158].

The causes of CLL are not well understood. One suggested risk factor is genetic predisposition: the familial incidence of CLL (defined as having one affected relative) is three times higher than that in the general population [164-166] and the relative risk for first-degree relatives is between three and eight [158,163,166,167]. Overall, a family history of leukaemia is reported in 5-10% of CLL patients [166,168,169]. CLL may also present earlier and in a more severe form in successive generations (although there is mixed evidence to support this effect [163,167,170-172]), and several genome-wide association studies have identified loci which may have increased susceptibility to CLL [158]. However, current UK guidelines do not recommend that family members are screened for CLL [158]. There is no good evidence that any of the environmental risk factors that are relevant in other forms of leukaemia (e.g. exposure to benzene or radiation) are relevant for CLL [20,163].

3.2.2 Presentation, diagnosis and staging

The median age of presentation for CLL patients is 65-70 years [158,159,163]. Patients may present with symptoms such as lymphadenopathy (enlarged lymph nodes), tiredness, night sweats, weight loss or symptoms of anaemia or infection [21,158]. However, 40-60% of patients are asymptomatic at the time of diagnosis [165], with 80% diagnosed as an unexpected finding following a routine blood test [158]. In terms of diagnostic investigations, blood smears are always taken and tested for lymphocytosis (an excess of a type of white blood cell), followed by immunophenotyping using flow cytometry to confirm

this finding [158,173]. Further investigations vary by clinical indication with bone marrow biopsy considered in uncertain cases or to provide prognostic information [158,173]. Computerised tomography scans are not recommended outside of clinical trials where they provide a baseline upon which response to therapy can be assessed objectively [21,158]. Serum markers such as CD23 and β_2 -microglobulin may predict (progression-free) survival but there is currently no standard assay to evaluate their levels [163,173]. Genetic information may also be useful at diagnosis. This is considered further in Section 3.2.3.

Disease staging may provide useful prognostic information, and one of two methods is generally used. The Rai approach divides patients into those with low-risk (Rai stage 0), intermediate-risk (Rai stage I or II) and high-risk disease (Rai stage III or IV), using information on haemoglobin levels and the presence of lymphocytosis and enlarged nodes [173]. The Binet approach divides patients into Stage A, B or C, based on the number of areas where there are enlarged lymph nodes, and other information (e.g. anaemia) [173]. Both staging systems predict outcomes in patients presenting with adverse clinical features, particularly when paired with other clinical and laboratory parameters such as age and gender (see Section 3.2.4 for more details). However, there is no evidence that information on staging should influence the timing or type of therapy [158].

3.2.3 Genetic testing in chronic lymphocytic leukaemia

Genetic testing in CLL indicates the presence of abnormalities which can inform prognosis (e.g. longer intervals between diagnosis and need for treatment) and influence treatment decisions (e.g. predict a poor response to fludarabine-based treatments). All of the abnormalities that are currently relevant in CLL (and which are discussed in this section)

are confined to tumour cells (i.e. somatic changes) and are not part of a patients' inherited genetic makeup (i.e. germline changes) [174]. Genetic defects related to CLL can therefore be acquired over time as tumour cells evolve, necessitating repeat genetic testing [173].

Clinically relevant genetic abnormalities in chronic lymphocytic leukaemia

The key genetic abnormalities in CLL are described in this sub-section, including 17p deletions and *TP53* mutations, 11q deletions, 13q deletions, trisomy 12 and Immunoglobulin Heavy Chain Variable (IgV_H) gene mutational status.

TP53 mutations and 17p deletions

Mutations in the *TP53* gene are associated with more advanced disease, poor response to treatment, and worse survival [21,23,158,159,163,173,175,176]. *TP53* mutations are not necessarily associated with 17p deletions (del 17p13) but the two are highly correlated: *TP53* mutations exist without 17p deletions in less than 5% of cases [23]. Together, these abnormalities are found in 3-37% of CLL patients [21,23,159,176-179], and their presence can influence treatment selection. Patients with these abnormalities do not respond well to treatments that kill CLL cells through a p53-dependent mechanism (e.g. fludarabine, cyclophosphamide and rituximab (FCR) treatment); outcomes are more favourable when treatments such as alemtuzumab are used [158,173,180].

11q deletions

Deletions in chromosome 11 (del 11q22) encompassing the *ATM* gene are found in 9-32% of patients [21,159,176-180] and have been associated with disease progression, poor

prognosis and short survival [21,158,175,176]. These patients may respond well to intensive combination therapy involving rituximab [158,176].

13q deletions

Two 13q deletions are relevant in CLL: del 13q14 (D13S319) and del 13q34 (*LAMP1*). The former is more common and is found in 50% of patients [176,181]. 13q deletions were originally thought to be entirely associated with more favourable outcomes (longer intervals between diagnosis and the need for treatment) [21,163,176,181]. However, evidence now suggests that patients with these deletions actually encompass two distinct sub-populations (Type I and Type II patients). Patients with only a 13q deletion (Type I) have a survival advantage [129], whereas Type II patients have a 13q deletion and increased genomic complexity (see following section), and thus a poorer prognosis [182].

Trisomy 12

Trisomy 12 is seen in 5-25% of patients [21,176-180] and represents a gain in genetic material, originally thought to be associated with advanced disease and shorter survival time [175,176]. Recent studies have, however, suggested that although progression-free survival (PFS) may be shorter in these patients, overall survival may actually be more favourable [176].

Other abnormalities with potential clinical relevance

Other abnormalities have potential clinical relevance in a small number of patients. Deletions in 6q are found in 2-9% of patients and may be associated with shorter time from diagnosis to need for first treatment and short survival [175-179,183]. 14q rearrangements

(or translocations) are found in more patients (33-54%) but their prognostic significance remains unclear [177,178]. Two abnormalities which have been documented in other cancers and which may ultimately be clinically relevant in CLL are 2p24 gains (3-10% of patients) [177,179] and deletions in 18q21 (7% of patients) [175,178].

Mutational status of IgV_H

CLL can be classified according to somatic IgV_H gene sequence homology into 'mutated' or 'unmutated' cases [129,173]. Half of all patients present with unmutated disease [159]. These patients may have inferior outcomes, with one study suggesting that median survival in mutated and unmutated patients is 25 years and eight years, respectively [129,158,163]. However, it is uncertain whether the expression of unmutated IgVH genes predicts response to treatment for individual patients [173].

Combinations of mutations

CLL patients often possess more than one genetic abnormality. It has been suggested that there are four primary abnormalities, which can be ranked in order of relative importance and relevance to treatment decisions: 17p deletions, followed by 11q deletions, trisomy 12, and 13q deletions [174]. Others have included IgV_H mutational status within rankings, dividing patients into three risk groups with implications for treatment [176]:

- Low risk patients with mutated IgV_H but no other mutations (20% of all patients);
- High risk patients with unmutated IgV_H or 11q deletions (70% of all patients);
- Ultra high risk patients with 17p deletions or *TP53* mutations (10% of all patients).

Genetic abnormalities of likely future interest in chronic lymphocytic leukaemia

The focus of genetics research in CLL has evolved in recent years, and there is now a growing interest in combinations of genetic changes (genomic complexity) and changes over time (clonal evolution). Future genomic tests in CLL will likely need to be sufficiently flexible to be able to provide information on these effects in addition to information on the abnormalities described previously in order to maximise their clinical utility.

Genomic complexity

Genomic complexity has been discussed in a variety of disease contexts but precise definitions are rare [184,185]. CLL is one exception. In this context, genomic complexity has been defined as the presence of three or more copy number alterations (CNAs) at least 20 kilo base pairs (kb) in size, and/or copy neutral loss of heterozygosity (cnLOH) regions at least two Mega base pairs (Mb) in size, in addition to the known CNA regions assayed routinely by fluorescent in situ hybridisation (FISH) testing (a test which uses fluorescent markers to highlight genetic abnormalities) [129]. CNAs (acquired alterations in one or more sections of DNA (deoxyribonucleic acid) which usually correspond to relatively large regions of the genome) occur in some (but not necessarily all) cancer cells, and do not reflect what is in the germline [24]. cnLOH is the term used when an individual carries two copies of a mutated, non-functional tumour suppressor gene. This mutation may have arisen *de novo* (for the first time) in the individual or may have been inherited.

Genomic complexity may be correlated with response to chemotherapy and poor prognosis, but cannot be assayed suitably using FISH approaches [129]. It is also linked to specific abnormalities: cases with 17p deletions have higher complexity than those with

only *TP53* mutations [23,182]. The clinical utility of information on genomic complexity depends on how it is defined. Genomic complexity is found in 21-25% of cases when defined as at least three CNAs [177], but is found in many more cases when more loosely defined. Few patients have no genomic complexity in their tumours [129].

Clonal evolution

Clonal evolution is the term given to the acquisition of new somatic genetic abnormalities over a patient's disease course, or increased mosaicism (greater proportion of leukaemic cells versus normal cells), or both. It is associated with an aggressive disease course and disease refractoriness, so is a useful prognostic marker that requires monitoring over time [186]. Clonal evolution can be detected using genomic testing, but, as with genomic complexity, agreement is yet to be reached on an exact definition [186].

Recommended genetic testing practice in chronic lymphocytic leukaemia in the UK

The British Committee for Standards in Haematology currently recommends that CLL patients undergo pre-treatment genetic screening for somatic *TP53* mutations. FISH testing is recommended, accompanied by Sanger sequencing in some hospitals (Section 3.3.1 describes Sanger sequencing in more detail). The measurement of biomarkers other than *TP53* is not recommended outside of clinical trials in patients for whom there is a clinical indication for therapy, due to the limited evidence base [158].

UK recommendations broadly match those in other countries. The European Research Initiative on CLL recommends that *TP53* mutation testing should be integrated into the diagnostic workup of CLL patients [187], while the European Society for Medical Oncology

recommends that FISH analysis is undertaken prior to the start of therapy. The utility of possessing information on IgV_H status is also noted, but testing is not recommended [159]. Guidelines produced by the National Cancer Institute in the USA make similar recommendations concerning FISH and IgV_H testing [173].

Health economic evidence to support recommended genetic testing practice in chronic lymphocytic leukaemia

As this thesis presents several health economic analyses of the use of genomic testing in CLL, a structured literature review was conducted to summarise the health economic evidence base which supports recommended genetic testing practice in CLL.

Literature review aims, objectives and methods

The aim of this review was to identify references which report any health economic evidence related to recommended genetic testing practice in CLL. All forms of testing with a genetic component in CLL patients were considered.

Literature searches were undertaken during October 2012 in Embase, Medline, the UK National Health Service economic evaluation database (NHS EED) and UK NHS Health Technology Assessment programme (both via the Centre for Reviews and Dissemination (CRD) database), and The Cochrane Library. Econlit was not used: a scoping search revealed no references to economics and leukaemia. The search strategies were designed to be inclusive (i.e. minimising false negatives); details are provided in **Appendix 7**. Keywords were identified by examining relevant references in the literature, previous related literature reviews, and Medical Subject Headings (MeSH). Searches were limited to full

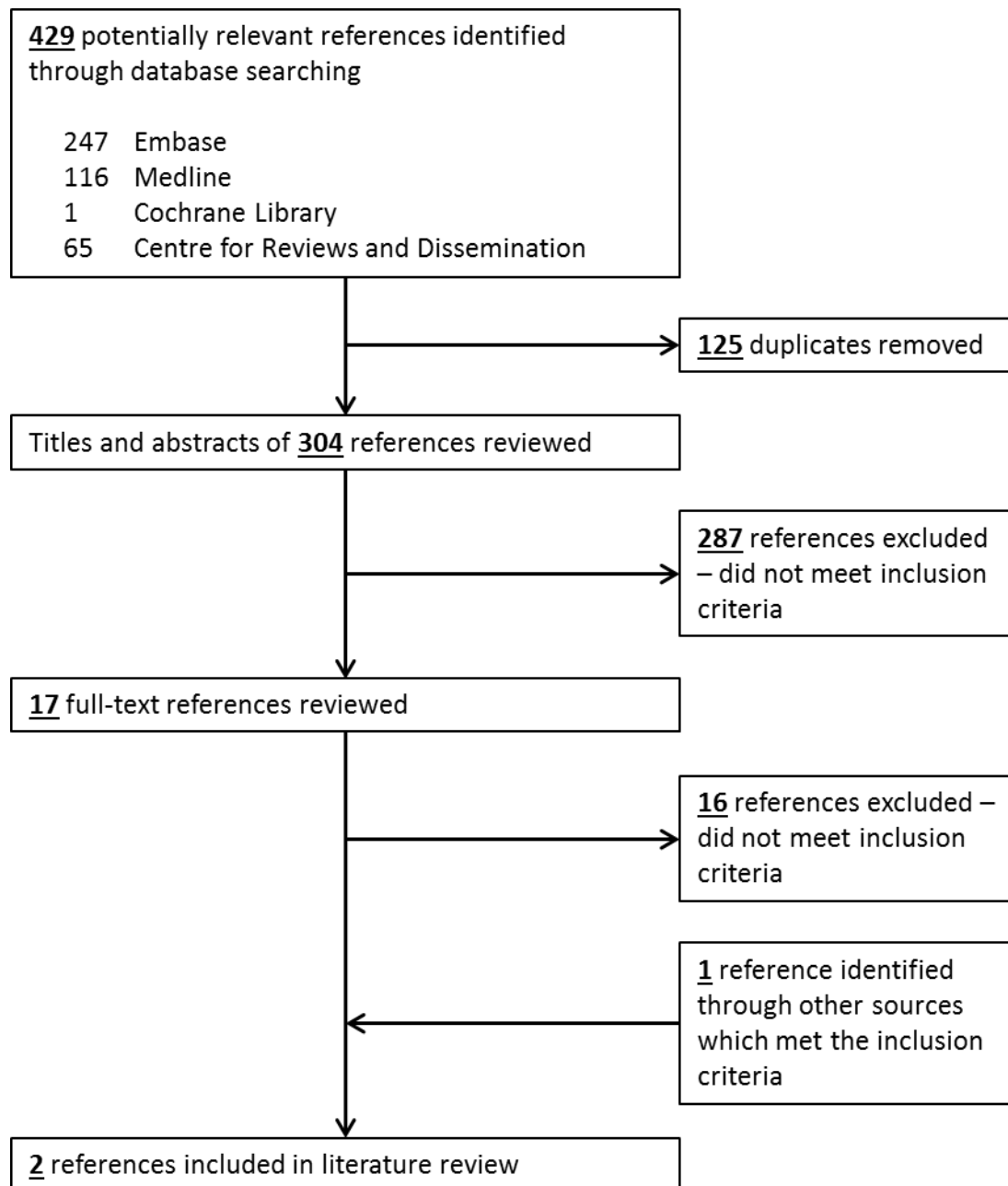
length references published in English (to enable in depth reviews to be carried out), studies in humans, and references published from 2000 onwards (to ensure that results were applicable to the present setting). Two national treatment guidelines for CLL in the UK have been published since 2000 which recommend a number of recently developed tests [158,163], hence health economic evidence in references published prior to 2000 is unlikely to be relevant, and has likely been superceded by more recent publications.

References which met the inclusion criteria were retrieved and health economic evidence related to recommended genetic testing practice in CLL was extracted. The references of reviewed articles were studied to identify potentially relevant studies, and all references which had cited the article in question were checked. In addition, the publication records of the authors of reviewed articles, and other authors known to be active in this field were examined. These publication records were obtained via PubMed searches and by reviewing author CVs (to identify publications not indexed by PubMed), where available.

Results

Figure 3.1 summarises the literature search results. 429 references were identified, with 304 remaining after removing duplicates. One reference met the inclusion criteria, and one was included from other sources, hence two references were included in the review. The health economic evidence base for genetic testing in CLL is therefore extremely limited. This is not a surprising result: the first exploratory studies of the use of FISH testing to inform patient management in CLL were only published in the 1990s [188-190], and evidence to support the widespread use of genetic tests to indicate prognosis in CLL was not widely available until the beginning of this century [163].

Figure 3.1: Literature review results – health economic evidence on genetic testing practice in CLL



The first study which met the inclusion criteria was an economic evaluation of a differential staining cytotoxicity (DiSC) assay, used to identify the sensitivity of CLL patients to different chemotherapies [191]. A cost-effectiveness analysis was conducted to compare this assay with current practice in the UK, with outcomes measured in life-years gained. The cost of

the assay was reported to be £380 (1998 prices). The average gain in life-years when the assay was used in patients over 65 years was 0.98 (men) and 0.64 (women). The cost per life-year gained with the assay ranged from £790 for male patients aged 50 with stage C disease to £6,410 for female patients aged 70 with stage A or B disease. The average incremental cost-effectiveness of the assay was estimated to be £1,470 per life-year gained. The authors concluded that the DiSC assay may be a cost-effective technology for improving CLL patient outcomes.

The second study which met the inclusion criteria investigated the use of multiplex ligation-dependent amplification (MLPA) analysis to identify genetic abnormalities in CLL patients, as compared to FISH testing [192]. This US study conducted micro-costing analyses for each test, with the total costs of MLPA and FISH testing estimated to be \$40 and \$234, respectively. MLPA was highly specific compared to FISH, with no false positive results, and had a detection sensitivity of 90%. False negative results were reported in seven patients. Cost and outcomes data were not combined by the authors, but they concluded that MLPA analysis could be used as a complementary method to FISH: although it is less sensitive, it is also less costly and allows simultaneous analyses to be conducted for more samples in a shorter time.

Neither of these tests are currently in widespread use for CLL patients in the UK. Given the limited nature of the health economic evidence base for genetic testing in CLL, only limited conclusions can be drawn from this literature review, namely that in specific patient populations, at specific locations, certain tests may be cost-effective (or cost-saving) in terms of indicating prognosis or response to treatment in CLL patients.

Current genetic testing practice in UK laboratories and hospitals

Although guidance exists which recommends the use of pre-treatment genetic testing in CLL patients, the limited evidence base underlying these tests suggests that usage may not be widespread. The reported availability of genetic testing for CLL does vary across the UK, with some form of testing offered in only 18 of the 74 laboratories (24%) providing genetic testing in the UK (personal communication, Idoia Gomez-Paramio, Orphanet UK and Ireland Project Manager, 17th April 2014). Most of these laboratories (listed in **Appendix 8**) offer testing to identify *TP53* and *ATM* deletions. Around half also offer testing to identify del 13q14, del 13q34 and trisomy 12. Testing to identify del 6q, 14q rearrangements and IgV_H is not offered as part of routine diagnostics. The lack of widespread IgV_H testing is predicted by the literature, which notes that this test is labour intensive, expensive and beyond the capabilities of many clinical laboratories. Furthermore, it is difficult to standardise good practice [173,193].

Availability of testing does not necessarily correlate with use; it could be the case that samples are sent from local laboratories to specialist regional laboratories for genetic testing for CLL. Unfortunately, information on the extent to which these tests are actually used in clinical practice is not currently available. To address this evidence gap, a survey of UK haematologists was conducted to collect information on current genetic and genomic testing practice in CLL in the UK NHS. A brief overview of this survey is now provided. Full details of the methods and results of this survey are presented in **Appendix 9**.

The two objectives of this survey were to summarise current genetic testing practice and to identify the opportunities and challenges associated with the translation of new

genomic tests into clinical practice in CLL. Potential survey respondents included UK haematologists with an interest or specialism in CLL, identified via existing UK CLL networks. Respondents were asked questions concerning current access to, and use of, genetic testing, and potential future use of genomic testing in this context. A full list of questions is provided in **Appendix 10**. The survey was conducted online from January to May 2014 using the LimeSurvey online survey application [194].

Seventeen haematologists completed the survey. The number of potential respondents in the networks that were accessed was unknown, hence a response rate could not be calculated. All respondents reported having access to a laboratory offering genetic testing for CLL patients, most worked in NHS Teaching Hospitals and the mean time since qualification as a haematologist was 15 years. Most respondents reported good access to FISH testing for the main abnormalities in CLL, although no respondents had access to FISH testing for all key abnormalities. Fifteen respondents (88%) reported that they request genetic tests as part of the clinical evaluation of CLL patients at diagnosis, and all reported that they request genetic tests as part of the clinical evaluation of CLL patients prior to starting treatment. Almost all respondents noted that the main benefit of possessing genetic information prior to starting treatment is to allow patients to be stratified by likely response to therapy. Respondents were also asked to consider the likely clinical utility associated with the use of genomic testing at diagnosis and prior to starting treatment. They indicated a stronger preference for genomic testing prior to treatment, rather than at diagnosis, and were more convinced of the clinical relevance of better identifying patients who would not respond to chemotherapy than of identifying patients with tumours exhibiting genomic complexity or clonal evolution.

These results suggest that genetic tests are in reasonably widespread use amongst haematologists specialising in CLL. Although survey respondents were primarily based in NHS Teaching Hospitals, these results are likely representative of clinical practice in the UK NHS as most CLL patients are managed and treated at cancer centres in such hospitals.

3.2.4 Clinical management and prognosis

This section describes how clinicians use the information provided by genetic and non-genetic diagnostic and prognostic tests to manage the treatment of CLL patients in the UK.

Disease course and initiation of treatment

CLL is an accumulative disorder, with leukaemic cells initially gathering in the bone marrow before spreading throughout the blood system to the lymph nodes, liver and spleen [21]. It generally follows a chronic relapsing course, although the natural course of illness can vary between indolent and aggressive disease [21,23,24,195]. Treatment of early stage disease does not prolong survival and is not indicated for Rai stage 0 or Binet stage A patients [163,173]. Regular monitoring (e.g. annual full blood counts (FBCs) and clinical examinations by haematologist oncologists for those with stable disease) is recommended in these patients (often called 'watch and wait' or 'active surveillance') [158,159]. First-line treatment such as FCR therapy is usually initiated in patients in Binet stage C or Rai stage III/IV disease, or in patients in Rai stage I/II or Binet stage B disease, provided that there is active disease and the patient is sufficiently fit to undergo chemotherapy [158,173]. Second-line treatment such as alemtuzumab or ofatumumab is usually initiated in patients with resistant disease, short time to disease progression, or 17p deletions [173]. Rate of progression is important when considering the appropriate timing of treatment [158].

Treatment options

The clinical pathways for CLL in the UK NHS are summarised in **Figure 3.2**. These pathways are based on current clinical practice outside of clinical trials at the Oxford University Hospitals (OUH) NHS Trust (the Trust in which the case study technology is being developed and implemented). The pathways were developed following interviews with Dr Anna Schuh (a consultant haematologist based in this Trust).

Due to the clinical heterogeneity of CLL and the old age of many patients, several clinical pathways exist. Patients are usually placed in a pathway based on their fitness to tolerate chemotherapy ('Go-go' versus 'Slow-go') and *TP53* status (in the case of 'Go-go' patients). Fitness to tolerate chemotherapy is based on factors such as age, susceptibility to infection and the presence of co-morbidities [158]. Maintenance therapy is not currently offered (outside of clinical trials) as the clinical benefit is uncertain [158].

'Slow-go' patients

These patients are unlikely to be able to tolerate aggressive chemotherapies. Standard first-line treatment is chlorambucil or bendamustine and rituximab (unless patients have access to clinical trials offering other exploratory treatments) [158]. Second-line treatment is often a further course of chlorambucil (if PFS is longer than two years). Other options include FCR or combination bendamustine and rituximab (BR) therapy, if this can be tolerated [158,196]. 'Slow-go' patients with refractory disease may respond to ofatumumab, but as the National Institute for Health and Care Excellence (NICE) have ruled that this is not cost-effective in these patients, funding for treatment must be sought via the Cancer Drugs Fund (CDF) [197].

'Go-go' patients

These patients are able to tolerate more aggressive chemotherapies and follow one of two pathways, depending on their *TP53* status. For patients without a *TP53* mutation, first-line treatment is usually combination FCR chemotherapy. This leads to improved response rates, PFS and quality of life compared with chlorambucil (the previously recommended first-line treatment) or fludarabine alone [21,196,198], and has been recommended by NICE [198]. One-fifth of patients will receive BR instead. If patients relapse and have not acquired a *TP53* mutation, second-line therapy is either FCR again (recommended by NICE [199]) or BR. Patients with refractory disease (or patients who have acquired a *TP53* mutation following first-line therapy) receive ofatumumab combination therapy, with 10% also undergoing allogeneic transplantation. Transplantation offers these patients the best opportunity of achieving long-term disease-free survival and the possibility of a cure. However, it is a high risk option that some cannot tolerate [158] and patients with refractory disease generally have a poor prognosis [158,196]. For patients with *TP53* mutations, FCR treatment is not recommended due to likely non-response (the complete response rate is <20% in these patients) [158,196,200]. Alemtuzumab combination therapy may benefit these patients [158,165,173,180,196], but is no longer offered in the UK. Instead, ofatumumab therapy is recommended, consolidated with allogeneic transplantation in 10% of patients [158]. Refractory patients are commonly offered a second course of ofatumumab therapy.

Variations in treatment pathways around the UK

Information on current clinical practice within the OUH NHS Trust was supplemented by information on current practice in other Trusts. CLL specialists were identified at three

hospitals in which current practice was anticipated to differ from that in the OUH NHS Trust: Royal Liverpool University Hospital, Royal Marsden Hospital, and The Christie NHS Foundation Trust. These clinicians were asked to consider how clinical practice varied in their setting, compared with that in the OUH NHS Trust. A small number of variations were identified. These are described in **Table 3.1**. As these primarily affect either ‘slow-go’ patients or a small proportion of ‘go-go’ patients, current practice in the OUH NHS Trust was assumed to be broadly representative of current practice elsewhere in the UK.

Table 3.1: Variations in chronic lymphocytic leukaemia treatment pathways in other UK NHS Trusts

Current practice within the OUH NHS Trust	Variations in other NHS Trusts
First-line treatment is FCR or BR (‘Go-go’ patients), or BR or chlorambucil (‘Slow-go’), depending on co-morbidities	‘Slow-go’ patients are only treated with chlorambucil
Outside clinical trials, refractory ‘Go-go’ patients with or without <i>TP53</i> mutations currently receive ofatumumab (via the CDF), followed by a bone marrow transplant in 10% of cases	‘Go-go’ and ‘Slow-go’ patients with refractory disease without <i>TP53</i> mutations may receive bendamustine and rituximab
Refractory ‘Slow-go’ patients with or without <i>TP53</i> mutations currently receive ofatumumab (via the CDF)	If ‘Slow-go’ patients have <i>TP53</i> mutations, they will receive alemtuzumab
If new <i>TP53</i> mutations emerge in ‘Go-go’ patients after first-line therapy, these patients are treated with ofatumumab (via the CDF), followed by a bone marrow transplant in 10% of cases	‘Go-go’ patients with new <i>TP53</i> mutations after first-line therapy may receive bendamustine and rituximab

BR = bendamustine and rituximab; CDF = Cancer Drugs Fund; FCR = rituximab, cyclophosphamide and fludarabine.

Future treatments

The range of treatments available in CLL is constantly expanding and clinical practice is likely to evolve considerably over the next decade. Three recently developed B-cell inhibitor therapies are likely to impact on patient management in the near future: ibrutinib, idelalisib, and ABT-199 [201-203]. All show promise in terms of progression-free and

overall survival, although given their high cost, these therapies may need to be targeted at likely responders using genetic information. This will require more widespread use of tests such as targeted next generation sequencing. As of March 2015, both ibrutinib and idelalisib had been approved by the CDF for use in combination with rituximab to treat relapsed CLL [204].

Treatment response

Various criteria exist to judge response to treatment [158] and assessment of response includes the evaluation of both blood and bone marrow [173]. Patients can fall into several categories post-treatment:

- Complete remission (CR);
- Partial remission (PR);
- Progressive disease;
- Stable disease (equivalent to non-response);
- Treatment failure i.e. anything other than CR or PR;
- Relapse (CR or PR achieved, but followed by disease progression);
- Minimal residual disease (MRD) negative status (<1 tumour cell in 10,000 normal cells).

Until the turn of the century, the aim of treatment (and the end point of most clinical trials in CLL) was to achieve complete remission. However, there have been considerable improvements in MRD detection in the last 20 years, such that achieving MRD-negative status is now a viable target. Furthermore, it has been suggested that MRD assessment can

be used to guide therapy, and routine MRD monitoring to indicate the need for pre-emptive treatment could become a feature of future clinical practice [205].

Treatment complications

Autoimmune complications are common in CLL, occurring in 10-20% of patients. Direct antiglobulin tests can be useful in predicting the development of these complications, and first-line treatment is usually prednisolone [158]. CLL patients also have an increased risk of infection due to their compromised immune function. Infective complications may account for half of all CLL-related deaths [21,158,173,196], with immune deficiency disease occurring in 20-70% of patients, particularly those with advanced disease and following immunosuppressive therapy. Risk factors include advanced stage and number of prior therapies [158]. Immunisation is less effective in these patients [165] and therapies such as FCR may lead to more infectious episodes, requiring more frequent monitoring of FBCs [196]. Immunoglobulin replacement therapy can be given to reduce the incidence of bacterial infections and chemo-immunotherapy combinations may be recommended for patients considered to be at higher risk of autoimmune complications [158].

Prognosis

Median survival following diagnosis is ten years [163], reduced to 2-3 years in patients with a *TP53* deletion [159,176]. Other factors indicating a poor prognosis include advanced stage at diagnosis, fast lymphocyte doubling time, raised serum markers, heightened CD38 expression and unmutated IgV_H gene status [158,163]. Male survival rates are lower than those for females [20], while younger patients are more likely to die of CLL-related causes [163].

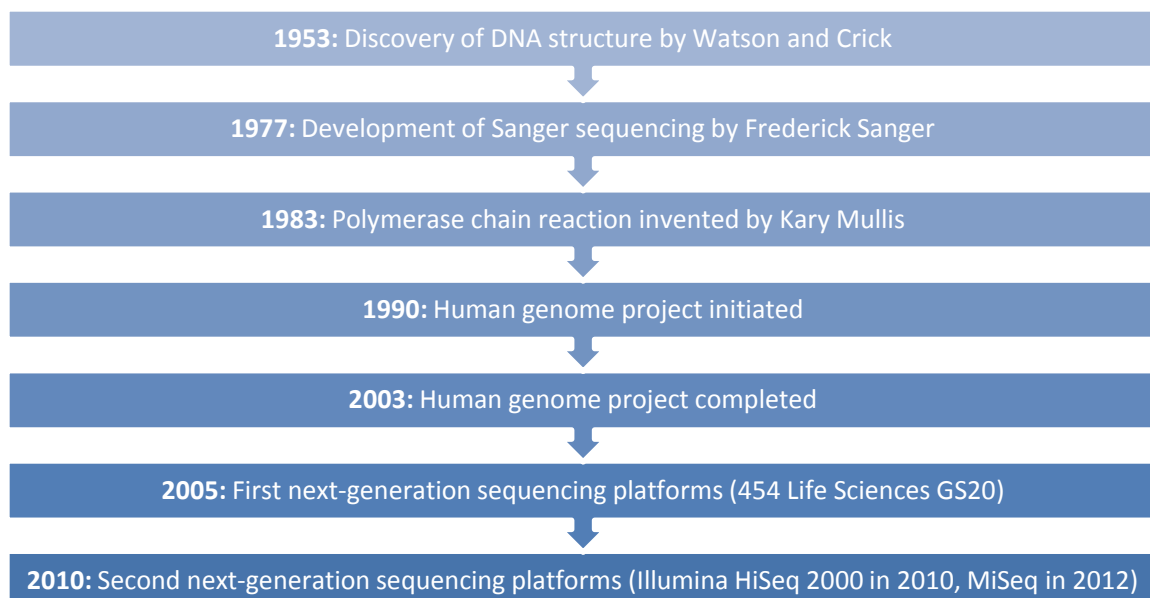
3.3 Targeted next generation sequencing

The first part of this chapter has provided an overview of current patient management and treatment practice in CLL, and described the rationale for using genomic testing in this disease context. This section will now describe the case study genomic testing technology evaluated in this thesis (targeted next generation sequencing) and summarise the economic evidence for the use of this technology in all disease areas.

3.3.1 Overview

Next generation sequencing (NGS) is a highly automated, high throughput and high read depth alternative to an older technique (Sanger sequencing), which remains a valuable, but time-consuming form of DNA sequencing. **Figure 3.3** presents a timeline of the evolution of this technique.

Figure 3.3: DNA sequencing timeline

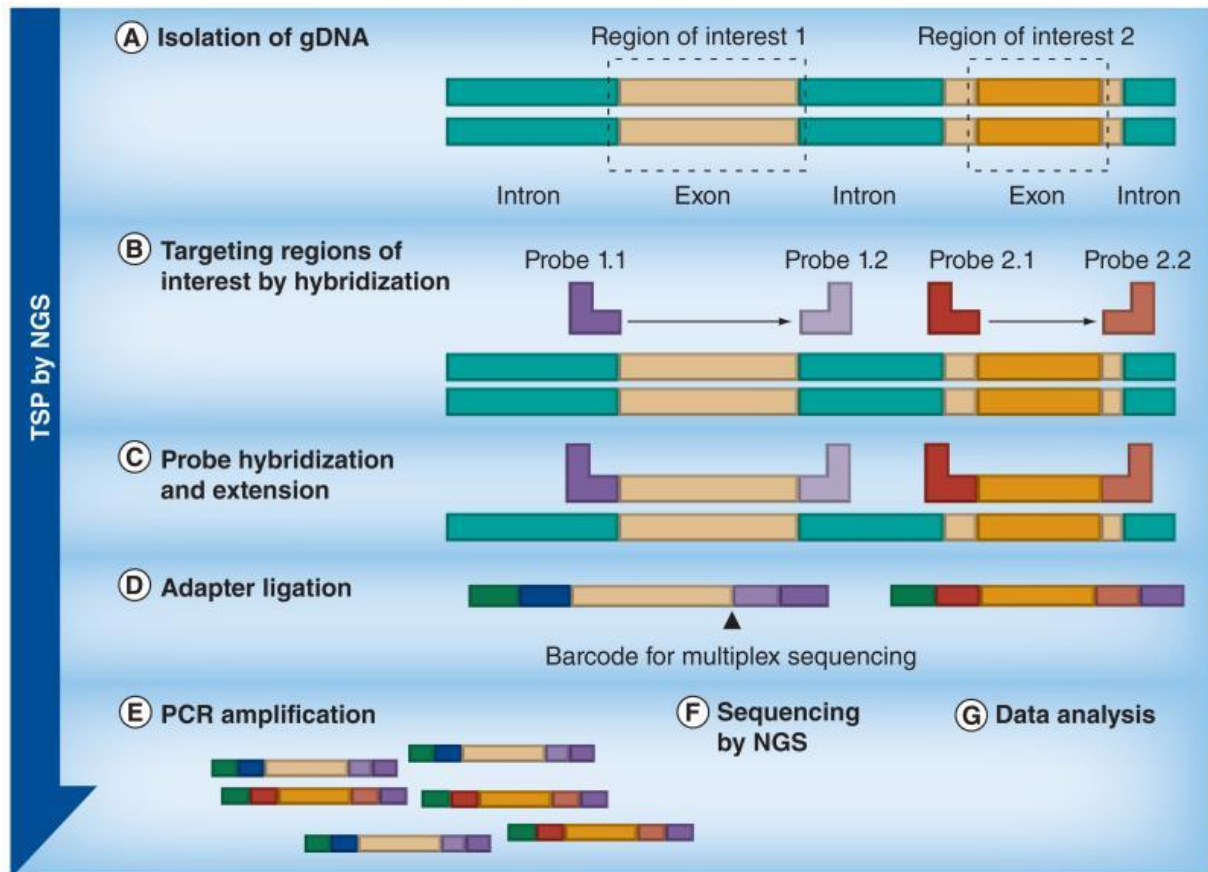


DNA sequencing is the process of identifying and determining the order of nucleotides within a DNA molecule. For cancer samples, this information can be compared with reference DNA sequences to identify mutations which contribute to tumorigenesis (the process by which normal cells transform into cancer cells) or influence response to treatment. The first rapid DNA sequencing methods were developed in the mid-1970's [206], with Sanger sequencing becoming the most widely-used approach in the 1980's. Sanger sequencing requires DNA to be extracted from samples, with the regions of interest then amplified using polymerase chain reaction (originally incorporating radioactively labelled nucleotides, but latterly incorporating fluorescently labelled nucleotides). Using DNA sequencing machines, DNA fragments are ordered by size (using gel electrophoresis), and the identity of each consecutive base is extrapolated to generate the DNA sequence.

NGS evolved from Sanger sequencing in the mid-2000's [207]. NGS instruments incorporate improvements in parallel processing and can sequence many DNA samples in a single batch, and multiple batches a day. NGS can take several forms. If the whole genome is sequenced including exons (protein-coding sequences) and introns (non protein-coding sequences), this is called whole-genome sequencing (WGS). If only exons are sequenced, this is whole-exome sequencing (WES). These approaches can sometimes miss low frequency clinically relevant mutations, which has led to the development of a third form: targeted NGS. This is a reduced form approach in which genes known to be involved in a particular disease are included in a targeted sequencing panel and can be sequenced in greater depth to minimise this error rate [208,209]. This approach allows more samples to be sequenced simultaneously than WGS or WES (each sequencing run produces a finite

amount of data, and targeted NGS produces less data per sample), often at a lower cost [210]. **Figure 3.4** summarises the targeted NGS process.

Figure 3.4: Overview of targeted next generation sequencing process



Reproduced from Alsolami *et al.* (2013) [209]

3.3.2 Targeted next generation sequencing in chronic lymphocytic leukaemia

Targeted NGS is a relatively recent technological advance, hence it is not currently used within current UK CLL clinical management pathways. Sanger sequencing is occasionally undertaken alongside FISH testing to identify *TP53* mutations in addition to 17p deletions, but only in the few laboratories offering this service. A key advantage that targeted NGS offers over Sanger sequencing in a CLL context is that it allows recurrently mutated genes to be tracked over time, aiding the detection of complex genomic changes such as clonal evolution [211,212]. Targeted NGS can also identify additional genetic changes relevant to

CLL by sequencing at higher or deeper coverage [213]. From a practical perspective, Sanger sequencing is slow, labour-intensive and not high-throughput [213], whereas targeted NGS can provide more information about the genetic abnormalities known to be associated with CLL and, compared to WES and WGS approaches, do so at a comparatively low cost, without the use of complex bioinformatics tools [209]. Overall, this means that targeted NGS may be able to identify more patients whose tumours are unlikely to maintain a response to standard first-line treatment (FCR chemotherapy) as compared to the current approach of FISH testing and/or Sanger sequencing. Avoiding FCR therapy in these patients (who represent around 25% of all CLL patients) spares them unnecessary treatment and associated side-effects.

One caveat to note is that targeted NGS may be surpassed by whole-genome approaches once these become price-competitive, so may not represent a long-term investment [209]. In addition, the evidence base for the use of targeted NGS in CLL is limited; those studies that do exist focus on whole-exome or whole-genome sequencing rather than targeted NGS [209,214].

3.3.3 Health economic evidence for the use of targeted next generation sequencing technology in all disease areas

As this thesis presents several health economic analyses of the use of targeted NGS in CLL, a literature review was conducted to summarise current economic evidence for the use of targeted NGS in all disease areas.

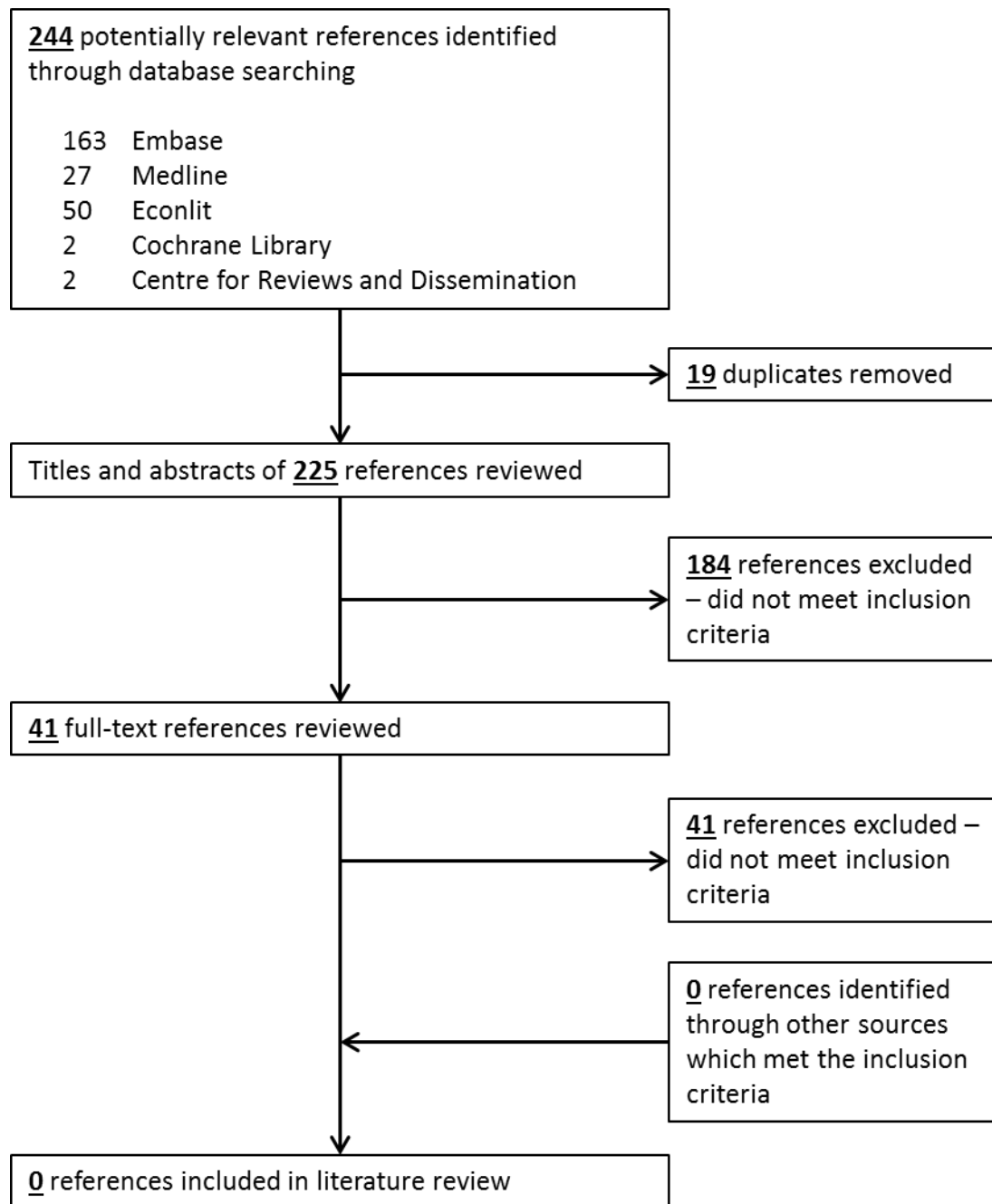
Literature review aims, objectives and methods

The aim of this review was to identify references which report the results of economic evaluations that consider the use of targeted NGS in any disease area. Although this search focused on targeted NGS, a variety of search terms were used to capture other forms of NGS, as different approaches are often mislabelled. Furthermore, the economic evaluation evidence base for these tests was expected to be small. Literature searches were undertaken in Embase, Medline, Econlit, NHS EED and UK NHS Health Technology Assessments (both via the CRD database), and The Cochrane Library. The search strategies used were again designed to be inclusive. Details of these strategies are provided in **Appendix 11**. The identification of keywords was carried out in the same way as in the literature review in Section 3.2.3, and the same search limitations were also applied. Searches were conducted in April 2014. References meeting the inclusion criteria were retrieved and health economic evidence related to targeted NGS was extracted. Reference lists, citation records and publication records were again carefully scrutinised.

Results

Figure 3.5 summarises the literature search results. 244 references were identified with 225 remaining after removing duplicates. None met the inclusion criteria. No references were identified from other sources. Given that targeted NGS technology represents a relatively recent technological advance, this was not a surprising result. Most references purporting to present data on the cost-effectiveness of this test instead made general statements about potential cost-effectiveness without conducting a full economic evaluation.

Figure 3.5: Literature review results - economic evidence for the use of targeted next generation sequencing in all disease areas



3.3.4 Barriers to more widespread use of targeted next generation sequencing technology

Targeted NGS technology is not currently in widespread use, either in CLL or in other disease areas. As this thesis considers the cost-effectiveness of implementing targeted NGS

in a new clinical area, it is important to note the barriers that have prevented the widespread application of this technology to date. In part, these barriers relate to a lack of clinical evidence (on test effectiveness). However, translational research evidence is also scarce, particularly economic evidence [11]. The literature review in the previous section clearly demonstrates this lack of evidence, while the literature review in Chapter Two which considered the methodological challenges associated with conducting economic evaluations in genomics highlighted specific health economic challenges in this context. There remains a perception that this form of testing is expensive compared to current approaches, although this issue may resolve itself over time if costs continue to fall at current rates [102]. There are also more practical challenges. In a survey of UK laboratories, respondents noted that genomic testing can be complex to implement, and that it requires significant investments of both time (e.g. staff training) and capital (e.g. equipment purchases) [215].

3.4 Discussion and conclusions

This chapter has provided an overview of CLL, describing the epidemiology of the disease, summarising current practice (including diagnosis and staging), and describing UK clinical pathways. There is some evidence that genetic testing could provide clinically relevant information in CLL, and genetic tests can be accessed by CLL specialists in several locations around the UK. However, there is little health economic evidence to support any form of genetic testing in this context. The focus of genetics research in CLL has evolved over time and information on both genomic complexity and clonal evolution is likely to be increasingly valued in the future, allowing clinicians to better target current first-line treatments.

This chapter also provided background information on targeted NGS, the case study genomic technology evaluated in this thesis. Targeted NGS can provide information on genetic mutations in CLL which could be used to inform treatment decisions, but the evidence base to support the more widespread use of this technology is currently limited.

Chapters Two and Three have provided an overview of the methodological, clinical and scientific background to this thesis. Chapters Four and Five will build on these foundations, describing the analyses that have been conducted in order to address the research questions set out in Chapter One. Chapter Four will begin by providing details of an extra-welfarist economic evaluation of the use of genomic testing in CLL.

Chapter Four | An extra-welfarist economic evaluation of genomic testing in chronic lymphocytic leukaemia

4.1 Introduction

Chapter Two described how the debate concerning the most appropriate economic evaluation approach has evolved, and reported the results of a literature review which suggested that the adoption decisions recommended by economic evaluation studies could vary depending on whether a welfarist or extra-welfarist approach was used. This chapter also considered the challenges associated with conducting economic evaluations of genomic technologies, noting that evidence generated using the welfarist approach might be valued by decision-makers in this context. Chapter Three then provided an overview of a specific application of genomic testing: the use of targeted next generation sequencing (NGS) to guide treatment decisions in chronic lymphocytic leukaemia (CLL).

Chapters Four and Five build on Chapters Two and Three by describing two analyses conducted to address the primary research question of this thesis: what is the most appropriate economic evaluation approach to use when assessing genomic interventions? Both the extra-welfarist (Chapter Four) and welfarist (Chapter Five) approaches are used to address the same decision problem, namely what are the costs and consequences associated with the use of genomic testing to guide chemotherapy treatment decisions in CLL patients in the UK? Broadly speaking, this entails a comparison between current practice (no genomic testing) and future practice (genomic testing). As current practice can vary (and is evolving constantly), and pre-treatment genomic testing could take several

forms, these analyses consider several comparators. More detail on these comparators is provided in Section 4.3.1.

This chapter presents the first of these analyses: an extra-welfarist economic evaluation of genomic testing in CLL. As noted in Chapter Two, this approach has grown to dominate decision-making in healthcare and has been adopted by all major decision-makers. Two different types of extra-welfarist economic evaluation were conducted: a cost-effectiveness analysis (CEA), with outcomes measured using life-years, and a cost-utility analysis (CUA), with outcomes measured using quality-adjusted life years (QALYs).

The chapter begins by describing the different ways in which the CEA and CUA could be conducted, before justifying the decision to use a decision analytic modelling approach. Sections 4.3 and 4.4 outline the process of building this decision model and populating it with transition probabilities, genetic parameters, unit costs and utilities. Section 4.5 describes the analytical methods, before Section 4.6 presents the model results. The chapter ends with a discussion of these results.

4.2 Selecting an analytical approach

All extra-welfarist economic evaluations require data on the costs and health-related outcomes associated with a healthcare intervention and at least one comparator (or a 'do nothing' alternative). As this data can be extracted from a variety of sources, data availability can drive the choice of analytical approach. Health economists often distinguish between two alternatives: a single study-based approach and a model-based approach [68]. In the case of the former, the economic evaluation is informed by one study (e.g. an

economic evaluation conducted alongside a randomised controlled trial (RCT)), from which data is extracted on resource use, health outcomes, intervention effectiveness and occasionally costs. With the latter, the same data is not available in a single study and must instead be extracted from several sources (e.g. multiple RCTs, observational studies, disease registries, unit cost databases) and synthesised using decision-analytic modelling.

For this economic evaluation, data on the use of targeted NGS to guide treatment decisions in CLL (compared to fluorescent in situ hybridisation (FISH) testing and Sanger sequencing) was provided by a retrospective study [24]. This study reanalysed patient deoxyribonucleic acid (DNA) samples from two national clinical trials to assess whether a targeted NGS approach could identify more of those patients who would not respond to usual first-line chemotherapy treatment (combination rituximab, cyclophosphamide and fludarabine (FCR)), compared to FISH and Sanger sequencing. This study therefore provided the data on intervention effectiveness that was required for this economic evaluation. However, at the time of this analysis, data on patient resource use, health outcomes and unit costs were not available from this or any other single study. Consequently, a model-based economic evaluation approach was used to synthesise this data from multiple sources.

The next step was to select a modelling framework. This should reflect the underlying process of the disease in question, capturing all of the main clinical and economic events. It is often the case that several modelling approaches could potentially meet this objective, ranging in complexity from simple decision trees to Markov models and discrete event simulations. Consequently, guidance is available on model selection [36,216-218]. To evaluate an intervention which can guide chemotherapy decisions in CLL, a model is

required in which individuals can be thought of as independent from each other (i.e. interaction is not important). The model must also be able to accommodate recurrent events (as CLL is a chronic cancer) albeit across a relatively small number of health states, and it should be possible to model the impact of a healthcare intervention (i.e. genomic testing) on the overall disease process. A Markov model is therefore an appropriate choice in this context. Markov models can be used to analyse uncertain yet repetitive processes over time (e.g. cycles of chemotherapy), are well suited to modelling long term costs and outcomes [36], and can be used to simulate intervention effectiveness by applying a treatment effect in different model states. Markov models are usually analysed at the cohort level: a group of patients enter the model and move through a variety of health states over a defined time period according to transition probabilities, accruing 'rewards' such as costs and health outcomes at the end of each model cycle. These rewards are then averaged across the whole cohort to estimate mean costs and outcomes. If this process is repeated for a healthcare intervention and at least one comparator, these costs and health outcomes can be compared in the form of an economic evaluation.

To validate this choice of approach, it is useful to review how previous model-based economic evaluations in CLL have been conducted. Five such analyses have been conducted as part of National Institute for Health and Care Excellence (NICE) health technology assessments (HTAs) in CLL, and critical commentaries on each of these studies have been prepared by evidence review groups (ERGs). Amongst other considerations, these commentaries judge whether the economic evaluations submitted by manufacturers have adopted the correct analytical approach (as specified in NICE guidance). This includes the selection of an appropriate modelling framework. **Appendix 12** summarises these HTAs

along with comments made by the ERGs about choice of modelling framework. All five studies used a Markov modelling approach (or a variation on Markov modelling) and all five ERG reports judged that this was appropriate, given the research question that was addressed. The approach taken in this thesis is therefore in line with previous modelling studies in CLL. The possibility of adapting one of these models for use in this study was considered, however none of these models were suitable for this purpose as they either failed to reflect current practice, did not consider the whole disease pathway or appropriate patient subgroups, or were unable to accommodate the genomic data required for this analysis due to a lack of complexity. A new Markov model was therefore constructed to conduct this extra-welfarist economic evaluation.

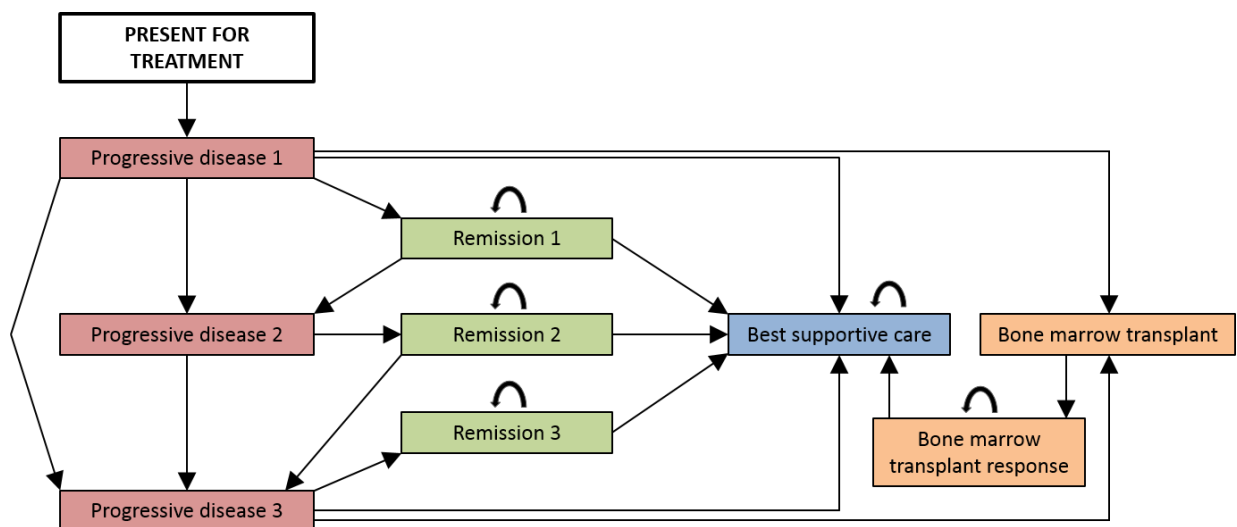
4.3 Devising the model structure

The Markov model was designed to evaluate the costs and outcomes associated with the use of genetic and genomic testing to guide treatment decisions in CLL. The first step when building this model was to consider the model structure, which should reflect all disease processes relevant to the affected patient population, and should capture all clinical and economic events that affect these patients. The starting point for this work was the UK NHS clinical pathways for CLL, presented in Chapter Three. These pathways split patients into 'go-go' patients (who can tolerate aggressive chemotherapy treatment, around 40% of all patients in the Oxford University Hospitals (OUH) NHS Trust) and 'slow-go' patients (the remaining 60%, who require less aggressive treatment). Genetic (or genomic) information is only likely to change the clinical management of 'go-go' patients at present. Furthermore, most of the samples that underwent genetic and genomic testing in the

retrospective study underlying this economic evaluation were taken from patients with these characteristics. This economic evaluation therefore focuses on this patient group.

The clinical pathways in the OUH NHS Trust, along with the models constructed within the NICE HTAs, were used to inform the development of a Markov model which simulated the disease pathway of ‘go-go’ patients. The structure of this model is shown in **Figure 4.1**. As per good practice recommendations, this does not directly map to a specific set of clinical guidelines [36]; rather, it is a generalised model which reflects the different disease states that can be experienced by this patient group, and is built in such a way that a variety of different versions of clinical practice can be incorporated into this framework.

Figure 4.1: A Markov model to simulate the disease pathway of ‘go-go’ patients



Patients enter the model when they present for treatment (it is unlikely that genetic or genomic testing would be carried out before this time). All patients undergo first-line treatment (Progressive disease 1), after which they cycle between remission (progression-free survival (PFS)) and progressive disease (symptomatic and receiving treatment) until

entering best supportive care. A small proportion of patients will undergo allogeneic transplantation (bone marrow transplant (BMT)), and then enter remission. Death from CLL-related causes is possible in all non-remission states, and death from other causes is possible in all states (transitions to death states are not indicated in **Figure 4.1** for brevity). These structural assumptions were validated by clinicians in the OUH NHS Trust.

4.3.1 Comparators considered in the model

The comparators considered in this economic evaluation cannot be narrowly defined as current practice (with genetic testing) and future practice (with genomic testing): pre-treatment genetic testing is not commonplace in all UK hospitals and there are several ways in which genomic testing could be implemented in CLL. The comparators must therefore include versions of current practice both with and without pre-treatment genetic testing. In addition, new therapies such as ibrutinib have been approved for use in relapsed/refractory CLL patients by the Cancer Drugs Fund (CDF) in England and are currently undergoing review by NICE. These treatments are likely to improve patient outcomes and will soon be added to clinical guidelines. This economic evaluation should therefore also reflect this likely change in clinical practice. Finally, although current evidence on ibrutinib treatment efficacy is focused on its use in relapsed/refractory patients [201,219], it is likely that in future evidence will emerge to also support its use in first-line treatment, particularly in patients who are unlikely to respond to FCR chemotherapy [220]. This potential scenario should also feature in the analysis.

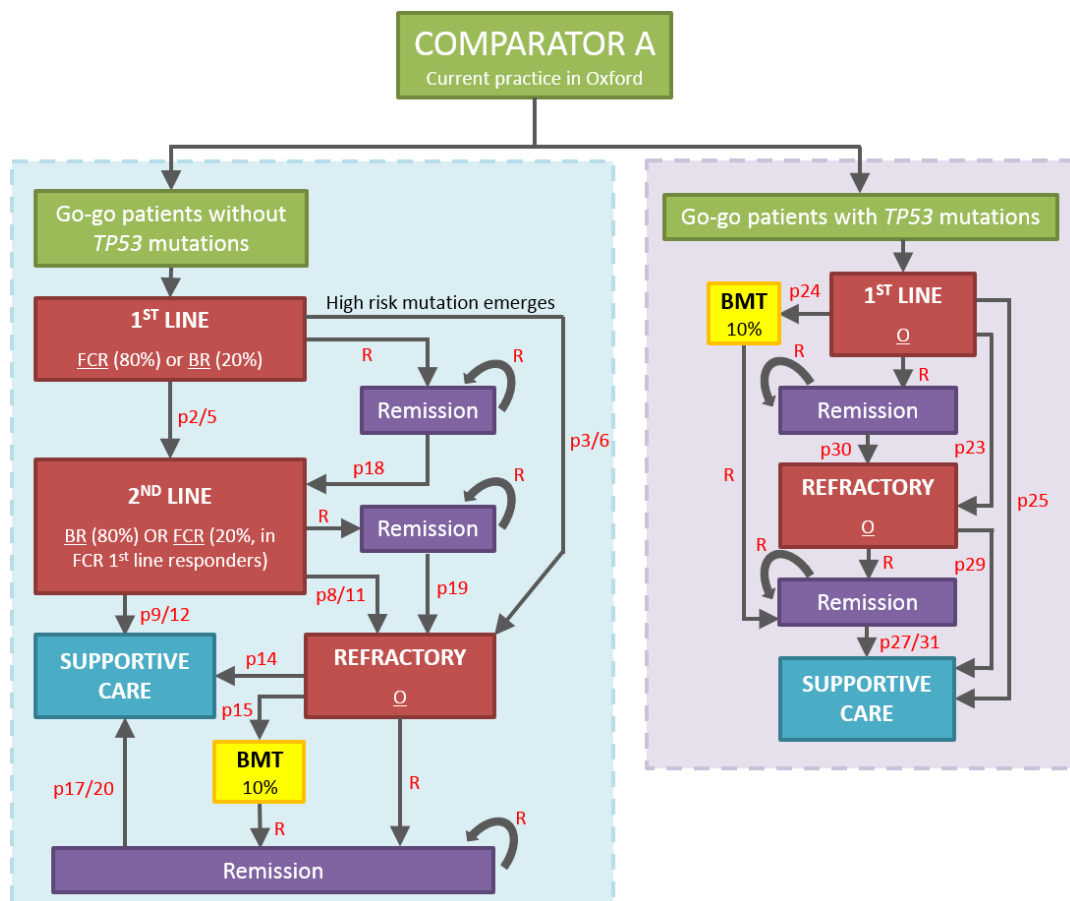
There are clearly a large number of potential scenarios that could be considered in this economic evaluation. Computational constraints and data limitations preclude the

inclusion of all potential comparators. Advice was sought from the CLL specialists consulted in Chapter Three as to which scenarios should be evaluated. Five comparators were selected, after considering the likelihood of each scenario: three current practice comparators, and two future practice comparators. Each of these strategies can be mapped on to the Markov model presented in **Figure 4.1**.

Comparator A - current practice

Comparator A (**Figure 4.2**) reflects current practice in the OUH NHS Trust.

Figure 4.2: Comparator A pathways



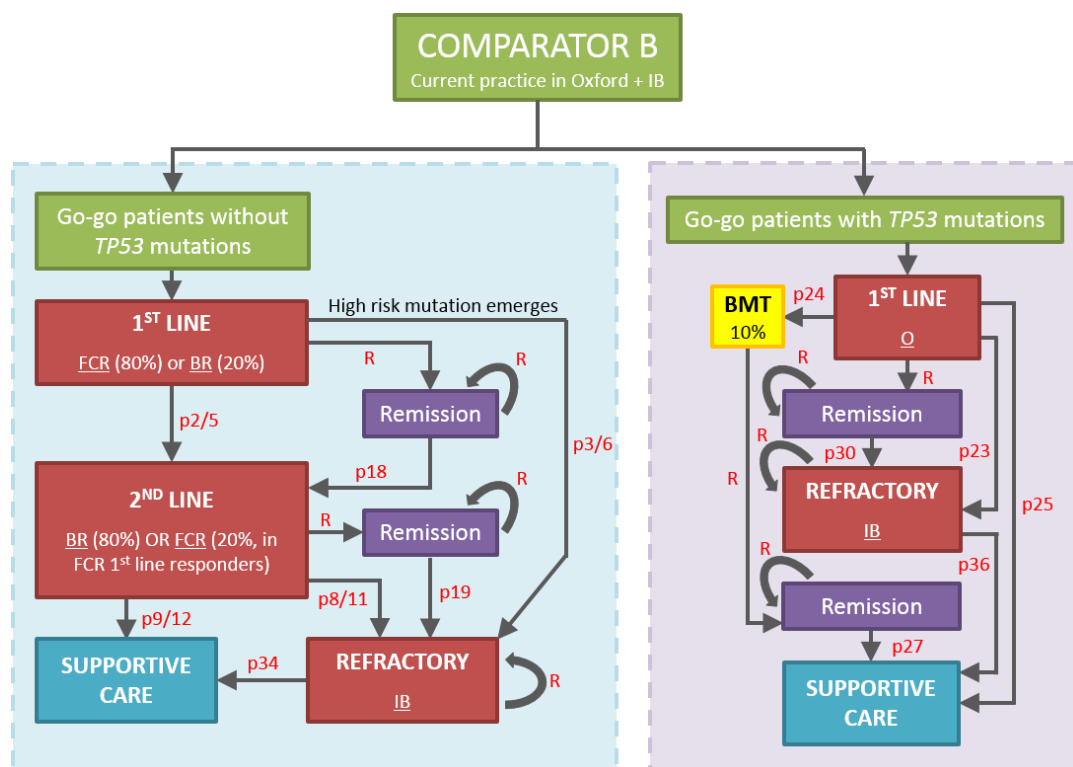
BMT = bone marrow transplant; BR = bendamustine and rituximab; FCR = rituximab, cyclophosphamide and fludarabine; O = ofatumumab.

This comparator is a sub-section of the pathways presented in Chapter Three, isolating the branches that are relevant for 'go-go' patients. The pathway begins with symptomatic patients undergoing genetic testing (FISH testing and Sanger sequencing) to identify those with *TP53* mutations. The two patient groups then follow different treatment paths.

Comparator B - current practice, plus ibrutinib

Comparator B (**Figure 4.3**) again reflects current practice in the OUH NHS Trust, now with ibrutinib given as refractory treatment for all patients. This is grouped with the other current practice comparators as genetic testing is still used to stratify patients by likely response to first-line FCR treatment. In all other respects, this matches Comparator A.

Figure 4.3: Comparator B pathways

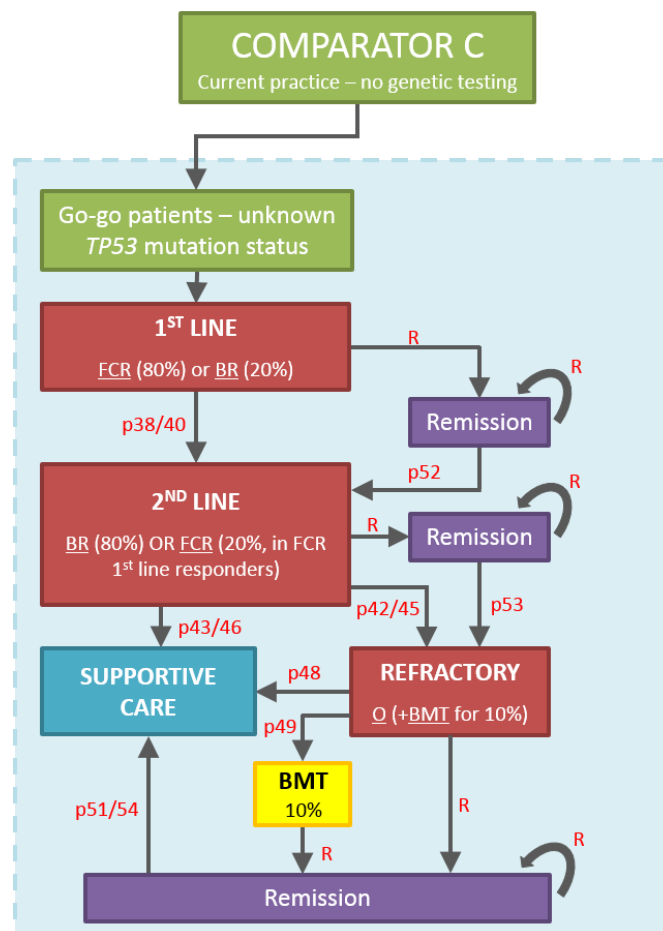


BMT = bone marrow transplant; BR = bendamustine and rituximab; FCR = rituximab, cyclophosphamide and fludarabine; IB = ibrutinib; O = ofatumumab.

Comparator C – current practice, no genetic testing

Comparator C (**Figure 4.4**) reflects current practice in hospitals which do not currently use genetic information to stratify patients by likely response to FCR treatment. This pathway is the same as that for patients with no *TP53* mutation in Comparator A.

Figure 4.4: Comparator C pathways



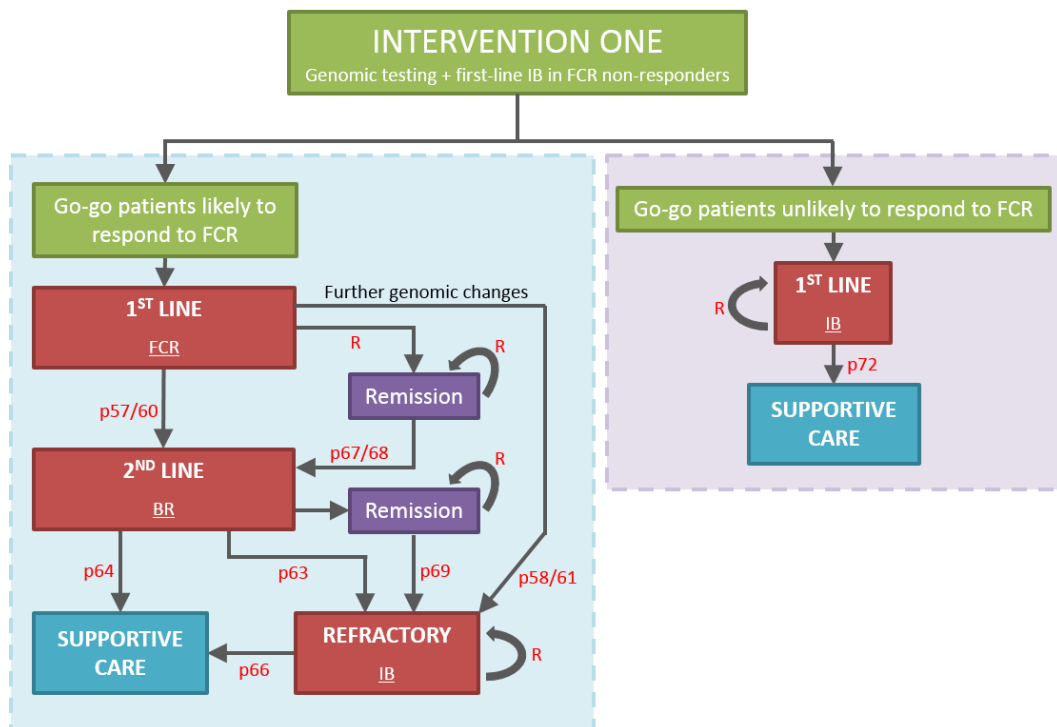
BMT = bone marrow transplant; BR = bendamustine and rituximab; FCR = rituximab, cyclophosphamide and fludarabine; O = ofatumumab.

Intervention One – future practice, with genomic testing and first-line ibrutinib treatment

Intervention One (**Figure 4.5**) reflects a possible configuration of future practice. Patients are stratified using genomic testing (targeted NGS) into those who are likely to respond to FCR treatment (FCR responders) and those who are unlikely to respond (FCR non-

responders). FCR responders follow a pathway similar to that for patients with *TP53* mutations in Comparator B. A key difference is that all patients predicted to be FCR responders receive FCR as first-line treatment and bendamustine and rituximab (BR) as second-line treatment (instead of the 80:20 split in current practice). FCR non-responders receive ibrutinib as first-line treatment until they no longer respond, then move into best supportive care.

Figure 4.5: Intervention One pathways

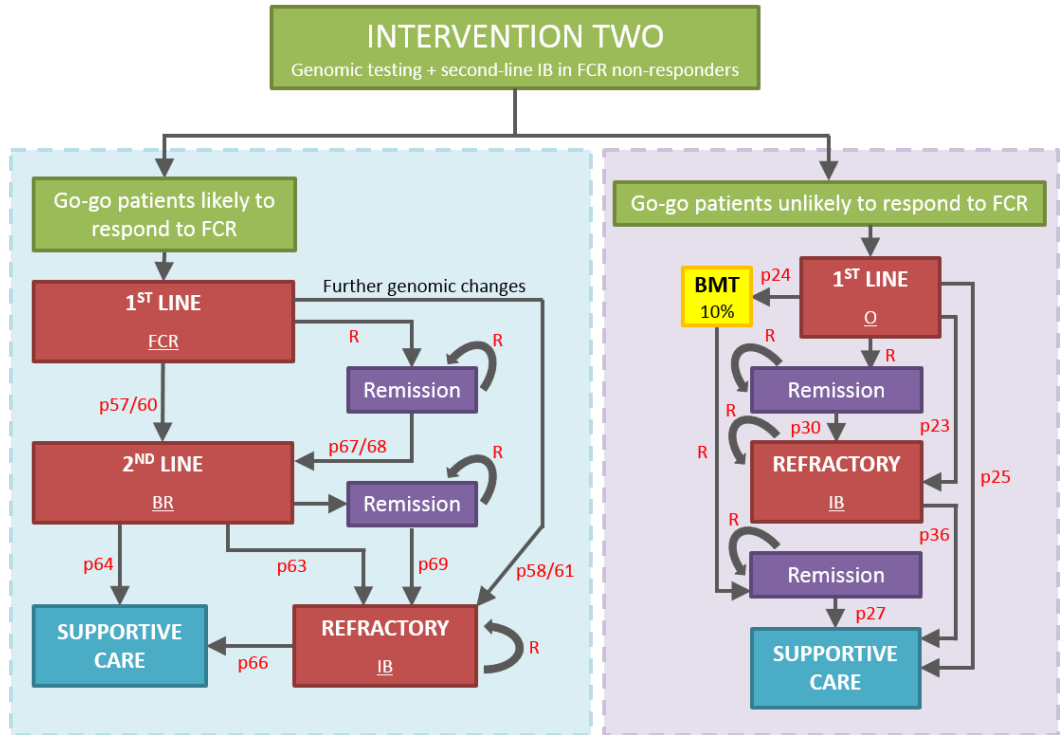


BR = bendamustine and rituximab; FCR = rituximab, cyclophosphamide and fludarabine; IB = ibrutinib.

Intervention Two – future practice, with genomic testing and refractory ibrutinib treatment

Intervention Two (**Figure 4.6**) reflects an alternative configuration of future practice. Again, patients are stratified using genomic testing, with FCR responders following the same pathway as in Intervention One. However, non-responders now receive ofatumumab as first-line treatment, followed by ibrutinib as refractory treatment.

Figure 4.6: Intervention Two pathways



BMT = bone marrow transplant; BR = bendamustine and rituximab; FCR = rituximab, cyclophosphamide and fludarabine; IB = ibrutinib; O = ofatumumab.

Table 4.1 summarises the key characteristics of the five comparators.

Table 4.1: A summary of the comparators considered in the Markov model

Comparator	Current or future practice?	Pre-treatment genetic or genomic testing?	Ibrutinib used?
A	Current	Genetic testing	No
B	Current	Genetic testing	Yes, as refractory treatment for all patients
C	Current	None	No
Intervention One	Future	Genomic testing	Yes, as refractory treatment for likely FCR responders and as first-line treatment for likely FCR non-responders
Intervention Two	Future	Genomic testing	Yes, as refractory treatment for all patients

FCR = rituximab, cyclophosphamide and fludarabine.

4.3.2 Additional design considerations

A number of additional design decisions must be taken before this model can be populated with transition probabilities, costs and outcomes.

Cycle length

The cycle length for the Markov model (i.e. the time that is assumed to pass before patients move from their current health state into the next health state) was set as 28 days (the usual length of time for one cycle of chemotherapy treatment). This is a common choice of cycle length for models in which patients undergo chemotherapy, as changes in the clinical management of such patients rarely occur more frequently than once a month [221,222].

Starting age of patients entering the model

The latest guidance from the British Committee for Standards in Haematology summarises the available clinical trial data for the initial treatment of 'go-go' patients, with median age of trial entry varying between 56-65 years [158]. As trial patients may, on average, be younger than non-trial patients, the upper bound of this age range (65 years) was applied in the base case (but varied in sensitivity analysis).

Time horizon

The total number of cycles over which the model is run is defined by the time horizon. Median survival in CLL is reported to be ten years, which acts as a lower bound for the time horizon [163]. The time horizons used in previous studies have ranged from ten years to 35 years [197,221-224]. As this analysis follows 65 year old 'go-go' patients from first-line treatment onwards, a time horizon of 30 years was used.

Time spent in chemotherapy treatment states

Transitions into and out of FCR, BR and ofatumumab treatment states are indicated in **Figures 4.2-4.6**, but for brevity the amount of time that patients can spend in these states is not displayed. As patients typically receive at most six 28-day cycles of each of these treatments (see Section 4.4.3 for further details), this restriction was applied in the model. Each FCR, BR or ofatumumab treatment state is represented in the model by six 28-day states. Patients must pass through all of these treatment states to reach a remission state, but can move from any of these six states to later lines of treatment (due to treatment non-response) or to death from any cause. A similar restriction is not applied to ibrutinib treatment states as ibrutinib is given continuously until patients no longer respond.

Bone marrow transplant state

Patients undergoing BMT enter a dedicated disease state which is modelled as a tunnel state: patients spend one cycle in this state, undergo a BMT, then enter a BMT remission state. The only other possible transition whilst in the BMT state is to death from any cause. This state is assumed to contain all of the costs related to the BMT procedure, while the BMT remission state contains all ongoing costs e.g. those related to complications.

More detailed summaries of all model health states are provided in **Appendix 13**.

4.4 Populating the model with transition probabilities, costs and outcomes

The parameters required to operationalise the Markov model include genetic parameters (related to the performance of genetic and genomic testing), transition probabilities (which govern how patients move between health states), and the cost and health outcome

‘rewards’ that are accrued when patients spend time in a health state. The process of estimating each of these parameters is described below.

4.4.1 Genetic parameters

Data on the efficacy of current genetic testing practice and future genomic testing practice was provided by a retrospective sample analysis study, introduced in Section 4.2 [24]. In this study, patient DNA samples underwent FISH testing and Sanger sequencing (current genetic testing practice) and targeted NGS (future genomic testing practice). The aim of both approaches was to identify genetic abnormalities such as *TP53* mutations, which have proven associations with prognosis and response to treatment in CLL. This study also provided evidence of new associations between abnormalities such as *SAMHD1* mutations and treatment outcomes [225]. The two testing approaches were compared in terms of the quantity and type of abnormalities detected, and these results were combined with data on PFS at 36 months in a series of multiple logistic regression models. These models predicted which patients would be PFS positive at 36 months (i.e. which patients had experienced disease progression). Model performance was assessed by comparing model predictions with actual data on PFS at 36 months.

This study has been described by Clifford *et al.* [226] and the results that are relevant to this economic evaluation are summarised in **Table 4.2**. The key results are contained in three regression models which reflect current genetic testing practice and two possible versions of future genomic testing practice. The current practice model used information on *TP53* mutation status from FISH testing and Sanger sequencing to predict that 7% of patients would be PFS positive at 36 months. Of these, 78% were true positives. Of the

residual group of 124 patients, 85 (69%) had not progressed at 36 months. Information on the proportion of patients in the residual and PFS positive groups was used in Comparators A and B in the Markov model in order to divide ‘go-go’ patients into those who did and did not have *TP53* mutations.

Table 4.2: Comparison between current genetic testing practice and future genomic testing practice

	Number of patients		
	Current genetic testing practice	Future genomic testing practice 1	Future genomic testing practice 2
	<i>TP53</i> only	<i>TP53</i> >5%	<i>TP53</i> >5% + <i>SAMHD1</i>
Predicted PFS positive	7% (⁹ / ₁₃₃)	11% (¹⁵ / ₁₃₂)	17% (²² / ₁₃₂)
True PFS positive (as % of predicted PFS positive)	78% (⁷ / ₉)	87% (¹³ / ₁₅)	82% (¹⁸ / ₂₂)
Residual group	93% (¹²⁴ / ₁₃₃)	89% (¹¹⁷ / ₁₃₂)	83% (¹¹⁰ / ₁₃₂)
Patients not progressing as % of residual group	69% (⁸⁵ / ₁₂₄)	72% (⁸⁴ / ₁₁₇)	75% (⁸² / ₁₁₀)
Total patients	133	132	132

PFS = progression-free survival.

The two models of future genomic testing practice both used information from targeted NGS on *TP53* mutation status to divide patients into likely FCR responders and non-responders. Specifically, these models considered patients with a *TP53* variant allele frequency (VAF) greater than 5%. VAF is a measure of the proportion of a specific gene variant (allele) among all allele copies being considered. This restriction meant that the future practice models focused on those samples in which *TP53* mutations were relatively more frequently observed (this is possible because targeted NGS provides additional information compared to FISH testing and Sanger sequencing). In addition, the second model of future genomic testing practice used information on the presence of *SAMHD1* mutations, another abnormality correlated with treatment outcomes.

In the first model of future practice, 11% of patients were predicted to be PFS positive at 36 months. The residual group – patients who would be considered to be FCR responders – contained 117 patients (89% of all patients). Of these patients, data on PFS at 36 months showed that 84/117 (72%) were actually in PFS at this time. The addition of information on *SAMHD1* mutation status in the second model of future practice resulted in fewer patients being classed as FCR responders (83%), but more of these patients (75%) were correctly identified. For this reason, this second model was used in the base case analysis to divide patients in Interventions One and Two into FCR responders and non-responders, and to divide FCR responders into those who were correctly and incorrectly identified as such. However, the use of the first model was evaluated in a sensitivity analysis.

4.4.2 Transition probabilities

The transition probabilities drive how patients move between states in the model and are primarily derived from published studies reporting the results of clinical trials of CLL treatments. Where possible, these probabilities were estimated using data extracted from Phase III studies. However, this Markov model considered a variety of treatments, in first-line, second-line and as refractory therapies, and for different genetic subgroups of patients. Phase III trial data were therefore not available to inform every transition probability required in the model, and in these cases Phase II and Phase II/Ib trial data were used. In all cases, the probabilities reported in these studies were converted into 28-day transition probabilities to match the model cycle length. No adjustments were made for gender differences as these were rarely reported. **Appendix 14** reports all transition probabilities used in the model, the data sources upon which these were based, and all calculations that were performed to derive these probabilities. This appendix describes in

detail the decisions that were taken in order to estimate the transition probabilities, however, some general comments on these decisions are noted below. Residual probabilities (R) are not reported in this appendix, but are indicated in **Figures 4.2-4.6**.

Lack of clinical trial data on ibrutinib treatment efficacy

In the model, ibrutinib is given as both first-line and refractory treatment in several patient subgroups. However, ibrutinib was only approved by the CDF in 2015 for use in relapsed/refractory CLL patients, based on data from clinical trials with small sample sizes, short follow-up periods and few patients in clinically relevant genetic subgroups [219,227-229]. Consequently there is considerable uncertainty around the transition probabilities that are informed by these data sources, particularly those relating to the use of ibrutinib as a first-line treatment, and in genetic subgroups. Some of these probabilities are based on averages of treatment effects from different trials, while others (e.g. the use of ibrutinib in FCR responders and non-responders) are extrapolated from data on treatment efficacy in other patient groups. As such, all of these probabilities were varied in sensitivity analysis.

Adjustments made to account for different genetic subgroups

A number of the comparators considered in this economic evaluation required data on treatment efficacy broken down by the presence or absence of genetic abnormalities (e.g. *TP53* mutations). For some of the more established treatments (e.g. FCR), this data was available [230]. For others, particularly BR and ibrutinib treatment, this data was lacking. Several approaches were used to fill this data gap. These included the application of hazard rates for treatment mortality from one study to data on treatment response from a second study (p8), and the use of an unweighted average of mortality rates from two separate

studies to inform the estimation of a transition probability for death from CLL for patients undergoing first-line ibrutinib treatment (p71). Again, all of these probabilities were varied in a sensitivity analysis.

Assumptions made regarding genomic testing

In Interventions One and Two, patients undergo genomic testing prior to beginning a new line of treatment. The transition probabilities applied in these scenarios reflect a number of factors. First, genomic testing is not perfectly sensitive or specific. Some patients will be incorrectly categorised as FCR responders while others will be incorrectly categorised as FCR non-responders. The proportion of patients in both categories is indicated in **Table 4.2**.

Second, the transition probabilities applied to these patient subgroups will differ. Patients correctly identified as FCR responders were assumed to transition out of first-line FCR treatment according to transition probabilities which reflected a 20% improvement on standard FCR treatment (in terms of mortality rate, non-response rate and PFS). This change in treatment efficacy was suggested by clinicians in the OUH NHS Trust and reflects the improved targeting of FCR treatment to likely responders. However, due to the limited evidence base underlying this assumption, this probability was varied between an assumed range of 10%-30% in the sensitivity analysis (see Section 4.5). Patients incorrectly identified as FCR responders did not benefit from this improved treatment effect, and instead transitioned out of first-line FCR treatment according to the transition probabilities associated with FCR first-line treatment in patients with *TP53* mutations. A similar improvement in outcomes in likely FCR responders receiving second-line BR treatment was also assumed. **Appendix 14** describes these adjustments in more detail.

In theory, similar adjustments should be made to the transition probabilities for patients identified as FCR non-responders. However, the consequences of misidentification for these patients are different. All patients identified as FCR non-responders (both correctly and incorrectly) are modelled as receiving ibrutinib (either as first- or second-line therapy). Hence, all of these patients receive a treatment which is likely more efficacious than FCR treatment (although some would have responded to FCR treatment, so any improvement in health outcomes is likely more modest in these patients). Moreover, the limited available evidence suggests that although there may be some minor variations in the treatment effect in this patient group by the presence (or absence) of specific genetic abnormalities, overall, ibrutinib still offers an improvement in treatment outcomes over and above all other treatment options [227]. Given this, no adjustments were made to account for the misidentification of FCR non-responders.

Death from other causes

Finally, to reflect the fact that patients in the model could die of causes other than CLL, age-adjusted UK life table data were used to estimate these probabilities [231].

4.4.3 Resource use and unit costs

Costs are accrued by patients when they use healthcare resources in any model state (except for states representing death, in which no resources are used). For the extra-welfarist economic evaluation presented in this chapter, a UK NHS perspective was adopted: only the costs incurred by the NHS when treating CLL patients were included. A number of cost categories are relevant for this analysis, including those associated with:

- Genetic and genomic testing;
- Clinical procedures, surgeries, hospitalisations and other diagnostic tests;
- Chemotherapy treatment (including drug costs, dispensing costs and delivery costs);
- Concomitant medications;
- Complications arising from treatment and clinical procedures (e.g. BMT);
- The time of clinical staff.

An overview of the decisions that were taken in order to attach unit costs from these categories to resource use in each model health state is provided in this section. Further details of all values, data sources and calculations are provided in the appendices indicated. It was noted in Chapter Two that economic evaluations of genomic tests may require a societal costing perspective to be adopted as such interventions may have impacts beyond the perspective of the healthcare sector (e.g. time lost from work due to poor health). An alternative analysis is presented in Chapter Six which applies such a perspective.

Estimating resource use in different health states

The healthcare resources used in each health state were estimated via interviews with Dr Anna Schuh (consultant haematologist) and other clinicians and specialist nurses in the OUH NHS Trust. **Appendix 15** details the resources that were costed in each health state. Many of the health states appear in more than one comparator. In most of these cases, resource use is similar. Any differences primarily relate to the use of either genetic or genomic testing, or varying complication rates in current and future practice.

Unit costs and data sources

The unit costs that were applied to resource use were extracted from several sources, including the British National Formulary, the UK Department of Health electronic market information database, the UK NHS Reference Costs database, and the Unit Costs of Health and Social Care report [232-235]. **Appendix 16** lists all unit costs used in this model, along with data sources. All costs are reported in British Pounds for the year 2013.

Genetic and genomic test costs

Two of the costs reported in **Appendix 16** (those for genetic and genomic testing) were not extracted from an external data source. Such costs are rarely available from NHS laboratories, and, as noted in Chapter Two, there are no national price lists for these tests. In addition, these are the unit costs for the two interventions that are compared in this economic evaluation. Consequently, it is important that these costs accurately reflect the resources consumed by the healthcare system when these tests are provided. Given this, a bottom-up microcosting analysis for both tests was conducted alongside the retrospective sample analysis study described in Section 4.4.1. This involved mapping current genetic and future genomic testing pathways, estimating the resources (equipment, consumables and staffing) used in both pathways, attaching unit costs to these resource use items, and calculating the total cost of both pathways [12]. **Appendix 17** describes this study in detail, including the methods, results and a sensitivity analysis.

Costing ibrutinib treatment

As ibrutinib has only been recently licensed for use in the UK, there is no information on the cost of treatment available in the public domain. There is also very little information

available on ibrutinib treatment costs in an international setting. Only one reference was identified, reporting a cost of \$273 for a daily dose of 420mg in a US setting [236]. This represents an annual cost of £63,011 for ibrutinib treatment (£4,830 per 28 day cycle), which was viewed by clinicians in the OUH NHS Trust as being unviable in a UK setting, given budget constraints and NICE cost-effectiveness thresholds. Given this, the annual cost in the UK was assumed to be £30,000 (£2,300 per 28 day cycle), which matches the reported upper limit of the NICE cost-effectiveness threshold [13]. Due to the uncertainty surrounding this estimate, this figure was varied within a wide range in a sensitivity analysis, and a scenario was evaluated which applied the higher costs from a US setting.

Complication costs

The costs associated with treating complications arising from both chemotherapy treatment and bone marrow transplants were also included in model health states, where appropriate.

Chemotherapy treatment

A variety of complications can be experienced by CLL patients following chemotherapy treatment. Grade one and two complications (e.g. chills) are not generally costed in economic evaluations in CLL as either no healthcare resources are used to treat these complications, or patients self-medicate [237]. Given this, only grade three and four complications were included in this analysis. These complication rates were estimated using the data sources which informed the transition probabilities, but were not estimated separately for different genetic subgroups as it is generally accepted that complication rates are independent of genetic subgroups. **Appendix 18** presents these complication

rates. The most common grade three and four complications (and the treatments given for these complications) were predicted by clinicians in the OUH NHS Trust, informed by clinical experience. These complications match those outlined in the latest UK CLL guidelines [158]. **Appendix 19** presents the costs associated with treating these complications. These unit costs were used to calculate an average cost for a grade three or four complication, which was combined with the data on average complication rates to calculate an average cost of complications for each health state.

Bone marrow transplant

The costs associated with complications experienced by patients after BMT were calculated in a similar way. **Appendix 18** and **Appendix 20** describe these calculations.

Estimating average length of treatment

Some of the resource use calculations described in **Appendix 15** required information on the average length of treatment for different types of chemotherapy. This data was extracted from the data sources which informed the transition probabilities and is presented in **Appendix 21**. Average length of treatment was not estimated separately for different genetic subgroups as this information was rarely available.

State costs

All of the information that was collected on unit costs and resource use was combined in order to estimate the overall cost of spending one cycle in each of the health states in the model. A summary of these state costs is provided in **Table 4.3**. A more detailed breakdown of the calculations underlying these state costs is presented in **Appendix 15**.

Table 4.3: Summary of the cost of spending a 28-day cycle in each model health state

State	Comparator	Cost per 28-day cycle
FCR treatment (first-line, first cycle)	A/B	£4,404
	C	£4,123
	Intervention 1/2	£3,958
FCR treatment (second-line, first cycle)	A/B	£5,373
	C	£5,092
FCR treatment (subsequent cycles)	A/B/C (first-line)	£4,289
	A/B/C (second-line)	£5,258
	Intervention 1/2	£3,805
BR treatment (first-line, first cycle)	A/B	£4,535
	C	£4,254
BR treatment (second-line, first cycle)	A/B	£5,542
	C	£5,261
	Intervention 1/2	£4,573
BR treatment (subsequent cycles)	A/B/C (first-line)	£4,427
	A/B/C (second-line)	£5,434
	Intervention 1/2	£4,427
Ofatumumab treatment (first-line, first cycle)	A/B	£12,798
	Intervention 2	£12,836
Ofatumumab treatment (first-line, second cycle)	A/B	£15,334
Ofatumumab treatment (first-line, subsequent cycles)	A/B	£4,455
Ofatumumab treatment (second-line or refractory, first cycle)	A/B	£13,054
	C	£12,773
Ofatumumab treatment (second-line or refractory, second cycle)	A/B/C	£15,590
Ofatumumab treatment (second-line or refractory, subsequent cycles)	A/B/C	£4,711
Ibrutinib treatment (first-line)	Intervention 1	£2,704
Ibrutinib treatment (second-line or refractory)	B	£2,780
	Intervention 1/2	£2,704
Remission following FCR treatment (all lines)	All	£212
Remission following BR treatment (all lines)	All	£130
Remission following ofatumumab treatment (all lines)	A/B/C/Intervention 2	£49
Undergoing BMT	A/B/C/Intervention 2	£43,724
Remission following BMT	A/B/C/Intervention 2	£233
BSC	All	£1,650

BMT = bone marrow transplant; BR = bendamustine and rituximab; BSC = best supportive care; FCR = fludarabine, cyclophosphamide and rituximab.

4.4.4 Health outcomes

The final set of parameters required to operationalise the model are the health outcome ‘rewards’ accrued by patients in each health state. As two forms of extra-welfarist economic evaluation were conducted (CEA and CUA), information was required on both life-years and QALYs. The number of life-years accrued by patients was estimated within the model, using information on the proportion of patients who were alive at the end of each cycle. This data was then combined with utility weights to estimate QALYs for each comparator.

Data on the utilities attached to health states experienced by CLL patients was extracted from the literature using a structured literature search. The search strategy and results are described in **Appendix 22**. In brief, an older literature search (conducted in 2010 [238]) was extended by applying the same search strategies for the period 2010 to 2014, with seven new studies identified. Utility data was extracted from all of these studies, and the most appropriate estimates were selected for use in the model. Sources which used preference-based approaches to estimate utility values for CLL health states using general population samples were favoured, as per NICE guidance [13]. The utility values that were selected were based on general health states (e.g. receiving a specific line of treatment, being in remission), and were not linked to specific treatments, as such data were not available in the literature. These utilities were then adjusted to reflect the complication rate in each health state (see **Appendix 18** for more information on complication rates), using information on the average disutility associated with a grade three or four adverse event. **Table 4.4** presents the utilities that were applied in the model.

Table 4.4: Utility weights used to calculate quality-adjusted life years in the model

Disease state	Utility	Notes	Source
Undergoing first-line treatment	0.803	No information was available on this utility in the literature. Beusterien <i>et al.</i> [239] report four utility values for patients undergoing first-line treatment, elicited from the general public using a standard gamble (SG) approach (complete response, partial response, stable disease and progressive disease). An unweighted average of these four values was used.	[239]
Undergoing second-line treatment	0.710	SG approach used to elicit preferences from general public.	[239]
Undergoing refractory treatment	0.650	SG approach used to elicit preferences from general public.	[239]
Undergoing BMT	0.650	No information was available on this utility in the literature. It was assumed that patients undergoing BMT would have the same utility as those undergoing refractory treatment.	[239]
In remission	0.910	SG approach used to elicit preferences from general public. This utility was also applied for patients receiving ibrutinib treatment (adjusted to reflect the complication rate) as these patients are assumed to be asymptomatic until they no longer respond.	[239]
Disutility associated with anaemia	-0.090	SG approach used to elicit preferences from general public.	[239]
Disutility associated with pyrexia	-0.110	SG approach used to elicit preferences from general public.	[239]
Disutility associated with pneumonia	-0.200	SG approach used to elicit preferences from general public.	[239]
Disutility associated with grade 3/4 adverse event	-0.133	Unweighted average calculated based on the disutilities associated with anaemia, pyrexia and pneumonia.	Calc
Receiving BSC	0.680	SG approach used to elicit preferences from general public. Napp <i>et al.</i> [238] adjust the Beusterien <i>et al.</i> [239] values to fit the decision context considered in a UK HTA.	[238,239]

BMT = bone marrow transplant; BSC = best supportive care; Calc = calculated; HTA = health technology assessment; SG = standard gamble.

4.5 Analytical methods

After constructing the Markov model and populating it with parameters, the next step was to conduct the economic evaluation. This section describes the analytical methods.

4.5.1 Base case analysis

Cohort analysis was used to conduct the economic evaluation, implemented using Microsoft Excel [240]. For each comparator, hypothetical cohorts of 10,000 patients were assigned to initial health states and their transitions between these states were modelled over time. As these transitions were assumed to occur at the end of each cycle, a half-cycle correction was applied [36]. The number of patients in each health state at the end of every cycle was observed, and the costs, life-years and QALYs accrued over time were calculated. A discount rate of 3.5% was applied to both costs and health outcomes, as per NICE guidance [13].

The mean costs and outcomes per patient for each comparator were calculated and compared in an incremental analysis. Comparators were ordered by increasing cost and incremental cost-effectiveness ratios (ICERs) were calculated (ratios of cost differences to outcome differences). Dominated comparators (i.e. those which were less effective than cheaper alternatives) and extendedly dominated comparators (i.e. strategies with an ICER that is higher than the next most effective strategy) were excluded in this incremental analysis. Strategies in the latter category produce additional gains in effectiveness at incremental costs higher than those of the next most effective strategy, so are dominated by a linear combination of two other strategies. Such strategies are not always excluded as by implication a proportion of the patient population will consequently receive a less

effective treatment. However, given that the use of genetic testing in CLL is not consistent across the UK, and this situation is expected to persist when genomic testing is introduced, this was considered to be a reasonable approach in this decision analysis. The ICERs of all comparators included in the final incremental analysis were compared to the standard NICE threshold of £30,000 per life-year/QALY gained in order to judge which would provide the most cost-effective use of limited NHS resources [13]. Finally, the CEA and CUA results were compared to determine whether there was concordance (both suggested the same adoption decision) or discordance.

4.5.2 Sensitivity and scenario analysis

Key model parameters were varied in univariate sensitivity analysis in order to assess the impact of parameter variations on the model results. These parameters included test and treatment costs, transition probabilities, complication rates, utilities and parameters such as the average number of cycles that patients received of each chemotherapy treatment. Parameter ranges reflected evidence from the original data source when this was available. In the absence of such evidence, parameters were varied by 50% above and below their base case value. Details of all parameter variations are provided in **Appendix 23**.

Several other parameters were varied in scenario analysis. These scenarios considered reductions in the cost of ibrutinib treatment (by 50%, 67% and 75%) as well as the use of US costs (£63,011 per year), no genetic and genomic testing before subsequent lines of treatment, and variations in the assumed improved treatment effects applied following genomic testing (from a 20% improvement in the base case to 40% and 60%). Different starting ages for patients in the model were also considered (25 and 85), as well as the use

of data from the alternative genomic testing model which excluded information on the presence of *SAMHD1* mutations. Finally, a scenario in which ibrutinib treatment was in widespread use was evaluated (i.e. Comparators A and C are no longer feasible strategies).

Probabilistic sensitivity analysis (PSA) was used to further consider the impact of parameter uncertainty within the model. PSA allows the uncertainty across all model parameters to be assessed simultaneously by assigning probability distributions to parameters based on measures of variance reported in the original studies [241]. Where parameters were informed by expert opinion, or no information on variance was available, a coefficient of variation equal to 0.10 was assumed to reflect this uncertainty [12]. Values from these distributions were randomly selected for each parameter and the model was run 1,000 times for different parameter combinations. The results of these simulations were used to calculate confidence intervals around the ICERs [36] and also cost-effectiveness acceptability curves (CEACs), which summarise the uncertainty surrounding the economic evaluation results by indicating which comparator is most likely to be cost-effective for different values of the cost-effectiveness threshold [242].

4.6 Results

This section presents the economic evaluation results. Prior to calculating these results, the model was debugged and the face validity of all parameter estimates was assessed (e.g. mortality rates were reviewed to ensure that they increased with subsequent lines of treatment, transition probabilities were checked to ensure that they fell between zero and one, and chance events were checked to ensure that probabilities summed to one). In addition, the number of patients in all health states in the model was reviewed every cycle

to ensure that this figure always summed to 10,000. Finally, model parameters were varied across a wide range to ensure that changes to the model results were intuitive.

4.6.1 Base case analysis

The base case results for the CEA and CUA are presented in **Table 4.5**.

Table 4.5: Base case results for the cost-effectiveness and cost-utility analyses

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.61	£71,576	£7,903	£7,903
	Int 2	6.65	£91,790	£580,390	EXT.DOM
	B	7.63	£107,703	£16,133	£35,376
	Int 1	7.45	£119,088	DOM	DOM
CUA	C	5.60	£69,704	-	-
	A	5.82	£71,576	£8,565	£8,565
	Int 2	5.93	£91,790	£177,198	EXT.DOM
	B	6.44	£107,703	£31,153	EXT.DOM
	Int 1	6.67	£119,088	£50,559	£55,891

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life-years; QALYs = quality-adjusted life years.

Cost-effectiveness analysis results

In the CEA, Comparator B (current practice with ibrutinib given as refractory treatment for all patients) yielded the most life-years per patient (7.63) while Comparator C (current practice with no genetic testing) yielded the least (6.37). Comparator C was also the cheapest strategy, costing £69,704 per patient, while Intervention One (genomic testing following by first-line ibrutinib treatment in likely FCR non-responders) was the most expensive, costing £119,088. With dominated strategies excluded, Comparator B was the most cost-effective strategy when compared to a threshold of £30,000 per life-year gained.

However, Intervention Two (genomic testing followed by refractory ibrutinib treatment in likely FCR non-responders) was extendedly dominated by a combination of Comparators A (current practice in the OUH NHS Trust) and B. When this comparator was excluded from the analysis, the ICER for Comparator B increased to £35,376, so the most cost-effective strategy was Comparator A, with an ICER of £7,093.

Cost-utility analysis results

The mean costs for each comparator were the same in both the CEA and CUA, so the cheapest and most expensive strategies were also the same. Comparator C also yielded the fewest QALYs (5.60) in this analysis, although Intervention One now became the most effective strategy, yielding 6.67 QALYs per patient. With dominated strategies excluded, Comparator A was the most cost-effective strategy, with an ICER of £8,565. This remained the case even when extendedly dominated strategies were excluded from the analysis.

Concordance of approaches

Both the CEA and CUA suggested that the most cost-effective strategy was Comparator A, with similar ICERs in both analyses. The five strategies were also ranked similarly in both analyses. The only difference was the strategy which yielded the greatest health gains (Comparator B in the CEA; Intervention One in the CUA). The use of ibrutinib as a refractory treatment in current practice (Comparator B) generates notable gains in life expectancy, but once these are weighted by the quality of life lived in these additional life-years, the use of ibrutinib as a first-line treatment following genomic testing (Intervention One) offers greater health gains.

4.6.2 Sensitivity and scenario analysis

Sensitivity analysis

The results of the univariate sensitivity analysis are presented in **Appendix 24**. As a large number of parameters were varied, this appendix only reports the results for parameter variations which changed the most cost-effective strategy from Comparator A to an alternative strategy, in either the CEA or the CUA. Only eleven parameter variations met this condition, suggesting that the base case results were robust to changes in single parameters. A summary of these results is provided in **Table 4.6**.

Of the variations that impacted on these results, only one changed the most cost-effective strategy from Comparator A to one of the intervention strategies. This was a reduction in the probability of moving to second-line treatment when in remission following first-line FCR treatment (effectively an increase in PFS), for FCR responders who had been correctly identified (see **Table 9a** in **Appendix 24** for further details). When a low value was considered for this parameter, the most cost-effective strategy changed from Comparator A to Intervention Two in both the CEA (£21,053/life-year gained) and the CUA (£19,248/QALY gained).

Table 4.6: Parameter variations which change the base case analysis results

Parameter variation	Low/high value	Most cost-effective strategy (ICER) ^a	
		CEA	CUA
Base case	-	Comparator A (£7,903)	Comparator A (£8,565)
Proportion of patients in Comparators A and B who do not have <i>TP53</i> mutations – varied between 85% and 95% (base case 93%)	Low	Comparator C (£10,943)	Comparator C (£12,447)
	High	Comparator A (£5,379)	Comparator A (£5,852)
28-day transition probability of dying from CLL during first-line FCR treatment: no <i>TP53</i> mutation – varied between 0.0014 and 0.0041 (base case 0.0027)	Low	Comparator A (£7,316)	Comparator A (£8,117)
	High	Comparator C (£5,855)	Comparator C (£6,921)
28-day transition probability of moving to second-line treatment when in remission following first-line FCR/BR treatment: no <i>TP53</i> mutation – varied between 0.0042 and 0.0126 (base case 0.0084)	Low	Comparator A (£8,566)	Comparator A (£9,626)
	High	Comparator C (£10,943)	Comparator C (£12,447)
28-day transition probability of moving to refractory treatment when in remission following second-line FCR/BR treatment: no <i>TP53</i> mutation – varied between 0.0092 and 0.0277 (base case 0.0185)	Low	Comparator A (£10,016)	Comparator A (£11,339)
	High	Comparator A (£29,590)	Comparator C (£12,447)
28-day transition probability of dying from CLL during first-line FCR treatment: any <i>TP53</i> mutation status – varied between 0.0018 and 0.0053 (base case 0.0035)	Low	Comparator C (£7,245)	Comparator C (£8,402)
	High	Comparator A (£7,671)	Comparator A (£8,529)

Parameter variation	Low/high value	Most cost-effective strategy (ICER) ^a	
		CEA	CUA
28-day transition probability of dying from CLL during second-line BR treatment: any <i>TP53</i> mutation status – varied between 0.0093 and 0.0280 (base case 0.0186)	Low	Comparator C (£25,590)	Comparator A (£12,298)
	High	Comparator A (£10,789)	Comparator A (£12,113)
28-day transition probability of moving to second-line treatment when in remission following first-line FCR/BR treatment: any <i>TP53</i> mutation status – varied between 0.0046 and 0.0139 (base case 0.0093)	Low	Comparator C (£8,520)	Comparator C (£9,586)
	High	Comparator A (£10,828)	Comparator A (£12,298)
28-day transition probability of moving to refractory treatment when in remission following second-line FCR/BR treatment: any <i>TP53</i> mutation status – varied between 0.0109 and 0.0327 (base case 0.0218)	Low	Comparator C (£10,049)	Comparator C (£11,392)
	High	Comparator A (£10,828)	Comparator A (£12,298)
28-day transition probability of moving to second-line treatment when in remission following first-line FCR treatment: FCR responder, correctly identified – varied between 0.0035 and 0.0105 (base case 0.0070)	Low	Intervention 2 (£21,053)	Intervention 2 (£19,248)
	High	Comparator A (£7,903)	Comparator A (£8,565)
Cost discount rate – varied between 0% and 5% (base case 3.5%)	Low	Comparator A (£9,698)	Comparator A (£10,511)
	High	Comparator B (£29,666)	Comparator A (£8,120)
Outcome discount rate – varied between 0% and 5% (base case 3.5%)	Low	Comparator B (£22,968)	Comparator A (£5,389)
	High	Comparator A (£9,580)	Comparator A (£10,355)

BR = bendamustine and rituximab; CEA = cost-effectiveness analysis; CLL = chronic lymphocytic leukaemia; CUA = cost-utility analysis; FCR = rituximab, cyclophosphamide and fludarabine; ICER – incremental cost-effectiveness ratio. ^a Parameter variations which change the base case analysis results are highlighted (Comparator B, Comparator C, Intervention 2).

Of the other ten parameter variations that affected the base case results, the most cost-effective strategy was always either Comparator B or C (i.e. a current practice comparator). Of note, the model was sensitive to the mortality rate for patients receiving first-line FCR treatment. When this rate was increased in patients without *TP53* mutations, the most cost-effective strategy changed to Comparator C in both the CEA (£5,855/life-year gained) and CUA (£6,921/QALY gained). This suggests that if the mortality rate for patients selected by genetic testing to receive FCR treatment becomes closer to that for patients with *TP53* mutations receiving first-line treatment, the benefit associated with genetic stratification disappears. Similarly, when the mortality rate was decreased in patients of unknown *TP53* mutation status (i.e. all patients receiving first-line treatment had better outcomes), the most cost-effective strategy again changed to Comparator C in both the CEA (£7,245/ life-year gained) and CUA (£8,402/QALY gained). Again, as the outcomes gap between the patient groups disappears, so does the benefit associated with genetic stratification.

The other parameters that impacted on the base case result included the proportion of patients identified as having no *TP53* mutation using current genetic tests, the PFS rate in patients without *TP53* mutations receiving first- and second-line treatment and the discount rates. Overall, for seven of the eleven parameter variations that affected the base case result, there was concordance in the results: the most cost-effective strategy changed in the same way in both the CEA and CUA.

Scenario analysis

Six alternative scenarios were considered: **Appendix 25** reports these results. The base case analysis results changed in four of these scenarios, as follows.

Scenario One: changes in the cost of ibrutinib treatment

All three reductions in the cost of ibrutinib changed the most cost-effective strategy to one which included ibrutinib. When the cost per cycle was reduced by 75% to £575, the most cost-effective strategy in both the CEA and CUA was Intervention One (genomic testing, then first-line ibrutinib treatment in likely FCR non-responders). Intervention Two (genomic testing, then refractory ibrutinib treatment in likely FCR non-responders) also became the cheapest strategy. A threshold analysis was also conducted to determine the ibrutinib cost at which one of the genomic testing strategies became the most cost-effective option. For the CEA, Intervention One became the most cost-effective strategy when the cost was £602 per cycle (approximately £7,848 per year). For the CUA, Intervention One became the most cost-effective strategy when the cost was £1,511 per cycle (approximately £19,830 per year). An increase in the cost of ibrutinib to match the cost of treatment in the USA (£63,011 per year) does not change the results, with Comparator A remaining the most cost-effective strategy in both the CEA and the CUA.

Scenario Three: variations in the improved treatment effect applied following genomic testing

An improvement in this treatment effect changes the most cost-effective strategy from Comparator A to Intervention Two in both the CEA and CUA. This is mainly due to the improvement in health outcomes rather than the smaller fall in health costs.

Scenario Four: variations in patient starting ages

Comparator A remains the most cost-effective strategy for patients aged 25. However, for patients aged 85, the most cost-effective strategy changes to Comparator C. This suggests

that when the underlying mortality rate increases, it becomes less important to stratify patients by response to treatment. However, few 'go-go' patients are likely to be this old.

Scenario Six: ibrutinib treatment in widespread use

In this scenario, ibrutinib was assumed to be in widespread use and Comparators A and B were no longer feasible strategies. As such, Comparator B was the most cost-effective option in the CEA while Intervention Two was the most cost-effective option in the CUA. This again indicates that the improved quality of life experienced by patients receiving ibrutinib treatment is an important driver of the economic evaluation results.

4.6.3 Probabilistic sensitivity analysis

The results of the PSA theoretically allowed 95% confidence intervals to be calculated for some of the ICERs presented in **Table 4.5**. However, the cost and effect pairs for several of these comparisons were spread across all four quadrants of the cost-effectiveness plane (see **Figures 1a** and **1b** in **Appendix 26**), suggesting considerable uncertainty. There was more certainty surrounding the ICER for Comparator B versus Comparator A in the CEA (**Figure 2** in **Appendix 26**), but considerable uncertainty surrounding the ICER for Intervention One versus Comparator A in the CUA (**Figure 3** in **Appendix 26**).

The results of the PSA also allowed CEACs to be calculated (**Figures 4.7** and **4.8**) which confirm that the most cost-effective strategy at a threshold of £30,000 per life-year/QALY gained is Comparator A. In both analyses, Comparator C is the most cost-effective strategy at low values of the threshold. The difference in life-years and QALYs between Comparators A and C is relatively small, and at lower values of the threshold there is only a small monetary value

attached to this difference, so the cost difference between the two strategies drives the CEAC results (with Comparator C being cheaper). The CEACs for the CEA indicate that there is no threshold value (up to £300,000) at which either of the genomic testing interventions are the most cost-effective option. The CEACs for the CUA suggest that Intervention One becomes the most cost-effective strategy at a threshold between £50,000-£55,000, which matches the ICER for the comparison between Intervention One and Comparator A.

Figure 4.7: Cost-effectiveness acceptability curves for the cost-effectiveness analysis

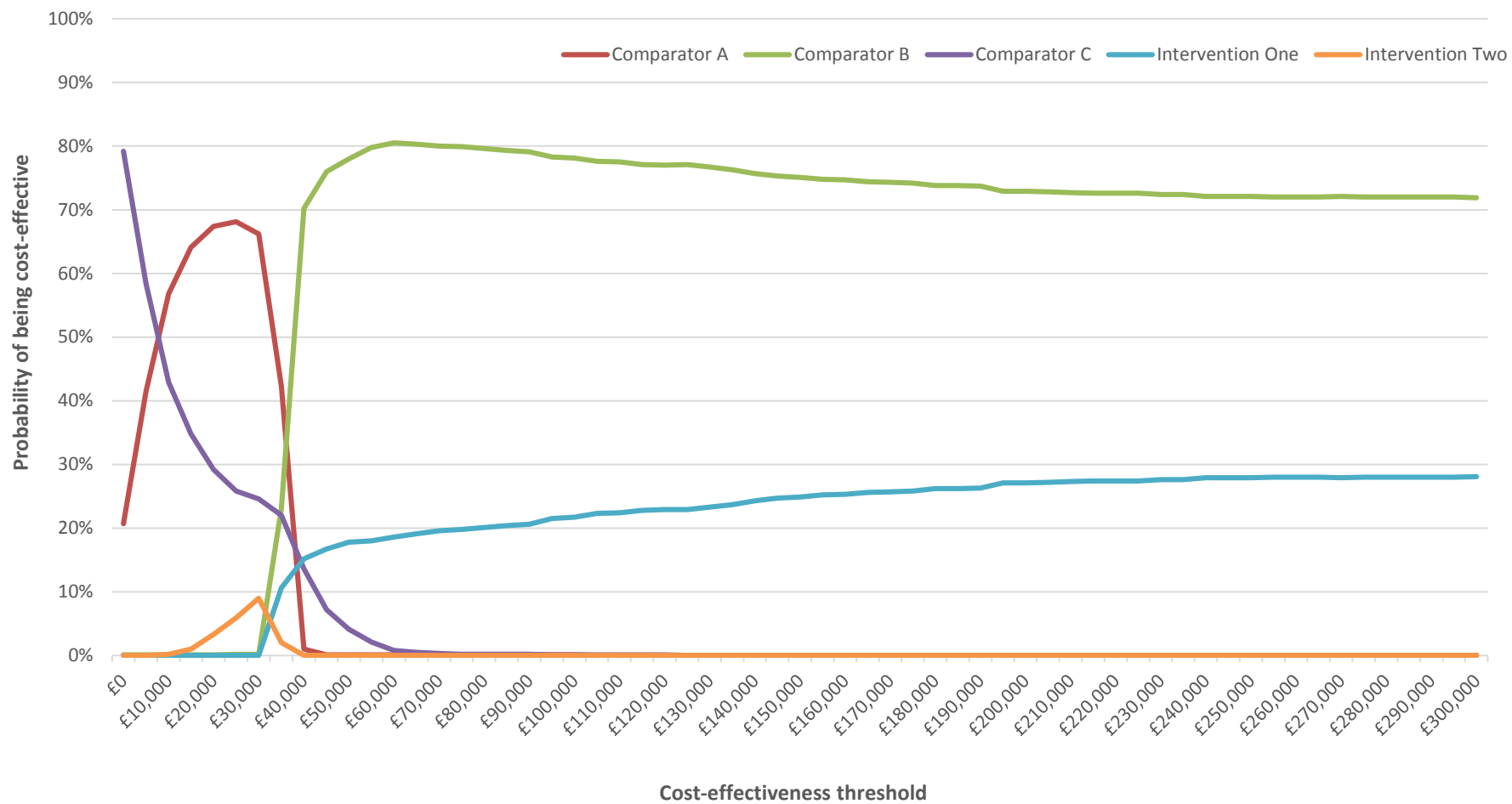
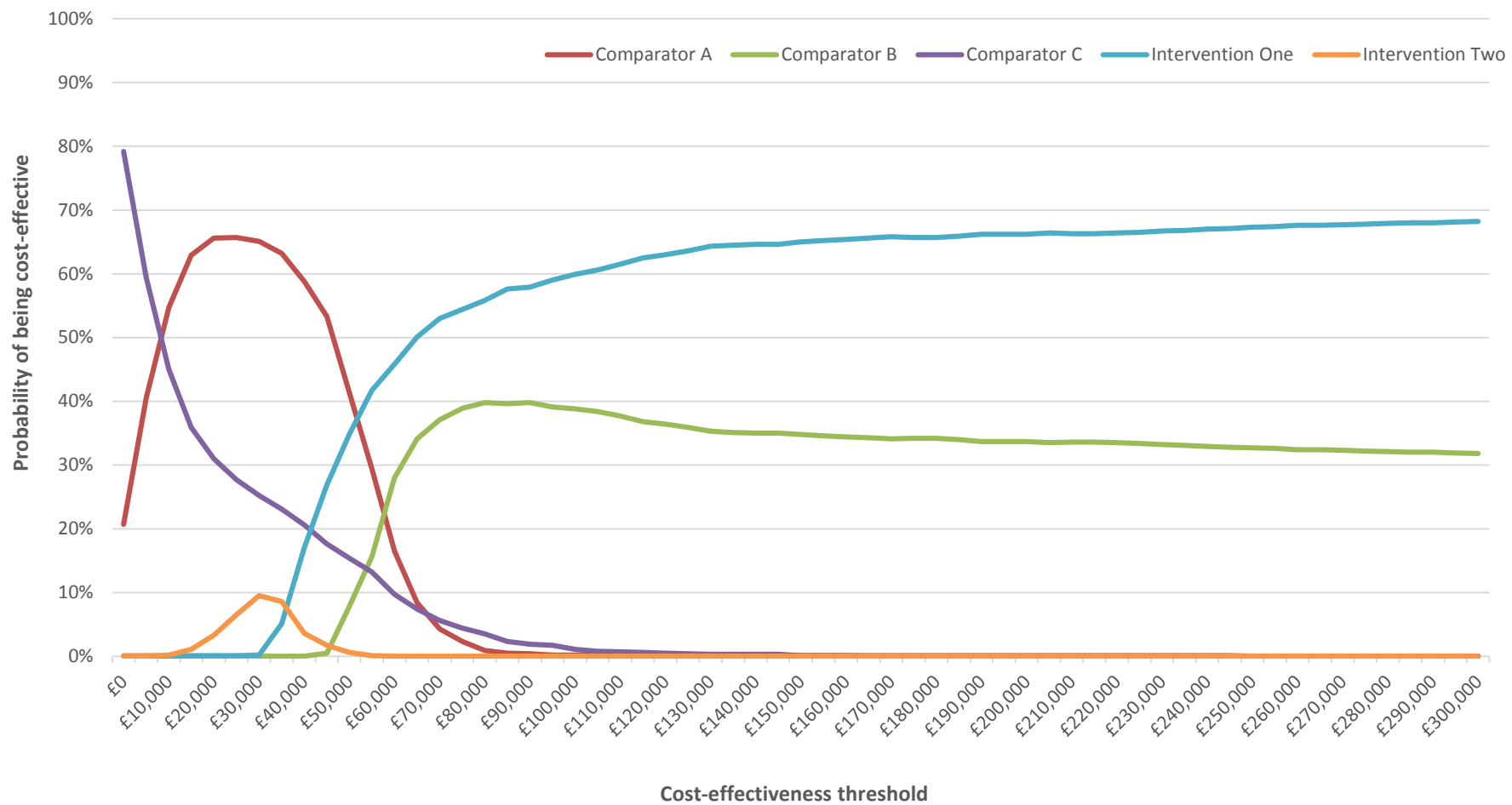


Figure 4.8: Cost-effectiveness acceptability curves for the cost-utility analysis



4.7 Discussion and conclusions

This chapter has described the development of a decision analytic model which was used to conduct two forms of extra-welfarist economic evaluation: a CEA which measured outcomes using life-years and a CUA which measured outcomes using QALYs. These economic evaluations considered five comparators, each of which represented a different way in which pre-treatment genetic or genomic testing could be operationalised for CLL patients in the UK.

The results from both the CEA and the CUA suggest that current practice in the OUH NHS Trust (Comparator A – genetic testing, no ibrutinib) is the most cost-effective strategy at a threshold of £30,000 per life-year/QALY gained. However, there is considerable uncertainty surrounding this result. Neither of the two genomic testing strategies are cost-effective in the CEA, while the cost-effectiveness threshold would have to increase to £50,000-£55,000 for Intervention One (genomic testing following by first-line ibrutinib treatment in likely FCR non-responders) to be the most cost-effective option in the CUA. The base case CEA and CUA results are generally concordant, and this concordance persists in the sensitivity analysis. The CEACs are, however, not concordant beyond the standard NICE cost-effectiveness threshold of £30,000 per life-year/QALY gained: at high values of this threshold, the CEA favours Comparator B while the CUA favours Intervention One.

The scenario analysis revealed that two key factors drive the cost-effectiveness of the genomic testing strategies. The first is the treatment effect applied following genomic testing: an improvement in this effect changes the most cost-effective strategy to Intervention Two (genomic testing, then refractory ibrutinib treatment in likely FCR non-

responders). This is important because this treatment effect is an assumption based on clinical expectations of the likely benefits of genomic testing. This suggests that generating better evidence on these effects should be a priority for future research.

The second factor that impacts on the cost-effectiveness of the two genomic testing strategies is the cost of ibrutinib treatment: a reduction in this cost leads to Intervention One (genomic testing followed by first-line ibrutinib treatment in likely FCR non-responders) becoming the most cost-effective strategy. This is noteworthy as the cost of ibrutinib treatment in the UK has not yet been announced by the drug manufacturer (Janssen Pharmaceuticals), hence the costs that are applied in this analysis are assumptions based on the anticipated behaviour of this manufacturer. This finding also suggests that if clinicians were able to specify treatment protocols that included shorter periods of ibrutinib therapy (possibly in combination with rituximab), strategies involving ibrutinib treatment and genomic testing may become the most cost-effective option.

The assumed treatment effect and ibrutinib cost are both limitations of this economic evaluation. Several additional limitations should also be noted. First, there is a lack of evidence on the effectiveness of CLL chemotherapy treatments; there are relatively few good quality Phase III clinical trials with large sample sizes. In those that do exist, there is limited data on treatment efficacy in the patient subgroups that are important for this analysis (e.g. patients with specific genetic abnormalities). This meant that several of the transition probabilities in this decision model were calculated by combining information from multiple data sources, which is not an optimal solution.

Second, data on the effectiveness of genetic and genomic testing was extracted from a retrospective sample analysis study which was undertaken to address the research question considered in this analysis [226]. However, the sample size of this study was relatively small; a larger sample would encourage more confidence in the economic evaluation results. In addition, this study was identified with the assistance of clinical collaborators, not by a literature review, which is not an optimal approach to evidence identification. However, this study represented the first genome-wide analysis of the genetic drivers of CLL using targeted NGS in an FCR-based first-line treatment setting. As such, it was highly likely to be the most appropriate source of evidence for this decision model.

Finally, although the utilities that were applied in the CUA were specific to CLL, they were not specific to different chemotherapy treatments, only to lines of treatment. This could mean that differences in utility by type of treatment were not captured in the analysis. However, these utilities were adjusted using information on complication rates and the disutility associated with complications, which may modify this limitation.

The strengths of the two extra-welfarist economic evaluations presented in this chapter should be acknowledged alongside these limitations. These economic evaluations represent two of the more comprehensive economic evaluations conducted in a CLL context to date. A variety of current and likely future clinical pathways have been evaluated, informed by the latest clinical trial data and reflecting a range of different treatments and treatment stratification strategies. The key parameters driving the economic evaluation results have been highlighted, and future research priorities have also

been identified. A more detailed discussion of how these strengths and limitations might impact on the selection of the most appropriate economic evaluation approach in this context will follow in Chapter Six.

To conclude, both extra-welfarist economic evaluations suggest that the current approach to treatment stratification for CLL patients in the OUH NHS Trust (using genetic testing) is more cost-effective than a new strategy which uses genomic testing instead. This is a result with notable implications for the allocation of limited healthcare resources in the UK, and also makes an important contribution to the CLL research agenda. The attention of this thesis will now turn towards Chapter Five and the second analysis conducted to address the same research question: a welfarist economic evaluation of genomic testing in CLL.

Chapter Five | A welfarist economic evaluation of genomic testing in chronic lymphocytic leukaemia

5.1 Introduction

The extra-welfarist approach to economic evaluation has been adopted by all major decision-makers in health care. However, it was noted in Chapter Two that in some circumstances, economic evidence generated using the welfarist approach might be valued by decision-makers in a genomic context, and may even impact on the adoption decisions recommended by economists. The welfarist approach has been applied relatively frequently to inform public sector decision-making in areas such as transport and the environment [243] (e.g. to evaluate the introduction of high-speed broadband [244] or the removal of a dam [245]). However, it is rarely applied in healthcare. A key barrier to more widespread use is that this approach – commonly operationalised in the form of cost-benefit analysis (CBA) – places a monetary value on human life, which can raise ethical concerns [86]. Chapter Two provides a more detailed exposition of the various issues surrounding CBA.

Given that the welfarist approach could provide useful information to decision-makers in genomics, this chapter presents the results of a CBA which evaluated the costs and consequences of using genetic and genomic testing to guide chemotherapy treatment decisions for chronic lymphocytic leukaemia (CLL) patients in the UK. This CBA required data on willingness-to-pay (WTP) for testing, which was collected using a discrete choice experiment (DCE). Although DCEs are a potentially informative source of data on WTP for genomic testing, operationalising this approach in a genomic context is not

straightforward, and there are several methodological obstacles preventing more widespread use. This DCE therefore also provided an opportunity to address some of these issues, including a comparison of visual and non-visual approaches to presenting genomic risk information to patients. Towards the end of the chapter these two strands of work are combined to consider the impact of different DCE design decisions on the CBA results.

The chapter begins by describing the process of selecting an analytical approach for the CBA, noting the advantages of undertaking a DCE and describing the issues surrounding risk presentation. Section 5.3 then outlines the steps taken when designing the DCE, before Section 5.4 presents the results and considers their implications. Section 5.5 describes how the CBA was structured then Section 5.6 presents the CBA results. The chapter ends with a discussion of the results of both the DCE and CBA, along with some concluding remarks.

5.2 Selecting an analytical approach

5.2.1 Potential analytical approaches

The basic analytical framework of a CBA is relatively straightforward: the costs and health outcomes (valued in monetary terms) of an intervention are compared directly, with a positive result indicating that an intervention generates a net benefit (i.e. the monetary value of intervention benefits is greater than the intervention costs). As the costing component of a CBA is very similar to that of a cost-effectiveness (CEA) or cost-utility analysis (CUA) the key analytical consideration relates to the monetisation of health outcomes, and several approaches exist to enable the calculation of WTP. These approaches fall into five categories, summarised overleaf.

The human capital approach

This approach uses market wage rates to value healthy time (i.e. time when an individual is not sick) and thus place a monetary value on an intervention by assessing the impact it may have on future earnings. Although it is relatively straightforward to operationalise this approach, there are several measurement difficulties related to the use of market wage rates (e.g. do these rates perfectly reflect the marginal productivity of workers?), and the need to value healthy time gained outside of the traditional workplace [12]. The use of the human capital approach is therefore problematic in the context of genomic testing in CLL for two reasons. First, it may not be possible to value the benefits of genomic testing (e.g. the value of possessing a correct diagnosis) using a measure which focuses only on the consequences for labour productivity. Second, many CLL patients are retired. Consequently, the application of the human capital approach either requires that the healthy time gained by these patients is ignored, or that it is valued using shadow prices (i.e. prices that are estimated rather than observed). Although approaches exist to generate shadow prices (e.g. the opportunity cost or proxy good approaches), there is no consensus amongst health economists as to the appropriate method to use [246].

The revealed preference approach

This is an indirect method of estimating WTP for health outcomes which examines the interaction between a non-marketed good and a market for another good [12]. An example of the use of this approach in a healthcare context is a study which considered the value placed on mobile mammographic screening units in rural areas of Australia. This study examined the travel costs incurred by women to use these facilities, with these costs viewed as a proxy for the value that women place on this intervention [247]. Travel costs

are often used to operationalise this approach, although other studies have also considered the relationship between wages and jobs which present health risks, using this information to calculate the value of a statistical life, which feeds into a CBA [12]. Neither approach is ideal in the context of genomic testing in CLL: testing is unlikely to require additional trips to hospital as it takes place during existing appointments, while a focus on labour market decisions may be problematic as few elderly patients will be working.

The stated preference approach: contingent valuation

The contingent valuation (CV) approach uses survey methods to present respondents with real or hypothetical scenarios in which they are asked to imagine that a market exists for an intervention, and then asked to reveal the maximum amount they would be willing to pay for this intervention in this market [12]. The CV approach has a long history in transport economics, environmental economics and market research, and the number of published CV studies in healthcare has grown in recent years [248]. Several methods can be used to elicit WTP values in a CV study, from relatively straightforward open-ended questions to more complex bidding games [86]. In general, whichever approach is taken, respondents are asked to indicate their WTP for the intervention as a whole. As a consequence, although the CV approach yields WTP information which can be used to inform a CBA, it provides limited information on which of the specific attributes of an intervention are valued most highly. Such information is likely to be useful to decision-makers in the context of genomic testing in CLL as it may well be the case that process-related outcomes are equally if not more important to patients than health-related outcomes.

The stated preference approach: discrete choice experiments

The CV approach is a relatively straightforward survey-based method for eliciting preferences for healthcare interventions. As the stated preference literature has evolved, more complex survey-based approaches have emerged in fields such as psychology and marketing, one of which is the DCE. DCEs are based on random utility theory (RUT), which proposes that individuals have a utility for all choice alternatives (e.g. different healthcare interventions, or different configurations of an intervention), which is made up of two components. The first is a systematic component, which relates to the observable attributes of the choice alternative (e.g. speed of test result). The second component is random (ensuring that choices are stochastic) and comprises the unidentifiable factors that affect choices (e.g. individual preferences). Consequently, RUT can be used to develop models which predict how choices will vary in response to changes in the attributes of each choice alternative, and thus obtain estimates of the utility associated with different healthcare interventions [249].

In practical terms, DCEs usually take the form of a series of choices in which at least two alternatives are specified in terms of their attributes, which can vary across a finite number of levels. Respondents are asked to complete these choice tasks in a survey and econometric techniques (generally a form of regression analysis) are used to analyse their responses. Models of choice behaviour are then estimated that yield coefficient values for each attribute and, if one of the attributes is a cost attribute, marginal rates of substitution can be calculated. These rates show how much respondents are willing to pay for unit changes in each attribute (e.g. WTP to reduce time to receive test results by one day), and this information can be combined to indicate WTP for the choice alternative as a whole.

Another attractive property of DCEs is that respondents are forced to make trade-offs between different attributes of an intervention. This makes it more likely that the preferences that are elicited are relatively unbiased, helping the DCE approach to avoid issues that are common to CV studies (e.g. the ‘warm glow’ effect that arises when individuals express support for a healthcare intervention) [86]. Furthermore, because the attributes of an intervention are valued individually in a DCE, this approach can be used to compare different ways of presenting attribute information to patients, and to test the impact of different presentational approaches on WTP. This is useful in a genomic context because some of the key attributes of genomic testing relate to risks, and previous studies have shown that respondents can struggle to interpret risk information as it is commonly presented in DCEs. This topic is discussed in more detail in Section 5.2.3.

DCEs are not the only survey-based method for eliciting preferences for healthcare interventions. One alternative is the best-worst scaling (BWS) approach, developed by Finn and Louviere in the 1990s [250]. This approach is related to DCEs, and is also based on RUT. BWS aims to overcome some of the limitations of DCEs by asking respondents to pick both the best and worst attribute levels in a choice set, thus obtaining more data from a given sample. BWS is potentially useful in this context, however, the first user guides have only appeared relatively recently [251]. Consequently this approach was not considered when deciding how to monetise health outcomes in this thesis.

Benefit transfer

The benefit transfer approach is common in environmental economics but rarely used in health economics. The aim of this approach is to estimate benefits in one context by

adapting estimates of benefits from another context. For this study, benefit transfer would require the identification of estimates of WTP for genomic testing to guide treatment decisions in a non-CLL context, then the adaptation of these estimates to reflect the operationalisation of genomic testing in CLL. Due to the limited application of this approach in healthcare to date, concerns surrounding the validity of the approach in general [86], and the lack of studies estimating WTP for targeted next generation sequencing in any context, this approach was not considered for the base case analysis in this study.

The analytical approach selected to estimate willingness-to-pay for genomic testing in chronic lymphocytic leukaemia

Given that the human capital, revealed preference, contingent valuation and benefit transfer approaches had several limitations in the context of genomic testing in CLL, the DCE approach was chosen to monetise health outcomes in this CBA. The following two sections provide further context for this decision. Section 5.2.2 begins by considering past applications of the DCE approach to inform CBAs (both in healthcare and other fields).

5.2.2 Existing cost-benefit analyses which use willingness-to-pay estimates generated by a discrete choice experiment

CBA guidelines often highlight the advantages (and disadvantages) of using DCEs to inform CBAs [252,253], and DCEs are regularly used as an input to CBAs in environmental and transport economics, particularly in the USA, Canada, Australia and the Netherlands. However, it is not straightforward to identify examples in the published literature as such studies are rarely intended for the public domain, and the DCE itself is often just a small part of the overall analysis (and thus not necessarily reported in detail, if at all). Examples

that do exist include a study in Australia which conducted a CBA to evaluate the costs and benefits of introducing nationwide high-speed broadband [244], and a study in the USA which reported the results of a CBA which investigated the costs and benefits of removing the Klamath Dam in Oregon and California [245].

Examples of CBAs informed by DCEs in healthcare are even rarer. A recently conducted literature review (personal communication, Rodolfo Hernandez, 9th June 2015) found just three examples, which are summarised below. Of these studies, two were identified via personal communications and the third was conducted by the author of the review.

McCormack *et al.* (2005) [25]

This UK health technology assessment considered the use of laparoscopic surgery for inguinal hernia repair compared to open mesh methods. An economic evaluation (CEA) was conducted using a Markov model which indicated that laparoscopic surgery was likely to be the most cost-effective strategy. A DCE was also conducted to elicit patient preferences for different hernia repair strategies. The results of this DCE were integrated into the economic model and indicated that laparoscopic surgery was unlikely to be associated with net benefits compared with open mesh methods.

Regier (2008) [254]

This study investigated the use of genetic testing to detect individuals with mental retardation, compared to usual testing. DCEs were conducted to collect data on WTP for true positives from testing in two groups: the general population and parents of children

with mental retardation. These WTP values were then applied as inputs to a decision tree model, which was used to conduct a CBA. The results of this CBA are yet to be published.

Burr *et al.* (2012) [26]

Burr *et al.* conducted both welfarist (CBA) and extra-welfarist (CEA and CUA) economic evaluations of different surveillance pathways for ocular hypertension, using a discrete event simulation model. In the extra-welfarist economic evaluations, the hospital-based surveillance strategy was the most cost-effective option, but at an incremental cost-effectiveness ratio greater than the standard NICE threshold. However, in the CBA, the same strategy was the only pathway to yield a net benefit.

In terms of common findings, it is pertinent to note that two of these studies used decision-modelling approaches that are similar to the Markov modelling approach applied in this thesis [25,26]. The same two studies also reported that the results of the welfarist and extra-welfarist economic evaluations were not concordant. However, given the limited evidence base, the scope for drawing out further common findings from these studies is limited. This chapter therefore moves on to consider a key challenge in operationalising the DCE approach in a genomics context: the presentation of risk information.

5.2.3 Issues surrounding the presentation of risk information in discrete choice experiments

The number of published healthcare DCEs has increased rapidly in recent years [255]. However, there are several challenges that must be addressed to ensure that the preferences elicited in these DCEs can be applied to usefully inform decisions in healthcare.

A key issue is that respondents are increasingly being asked to evaluate complex risk information (including probabilities, frequencies and percentages), particularly in DCEs evaluating new screening or diagnostic interventions [256]. Examples in genomics include a DCE evaluating preferences for pharmacogenetic testing to guide warfarin treatment [257], and a DCE evaluating preferences for the use of genomic testing to identify children with developmental disability [258]. This is problematic because the general public struggle with concepts of probability and uncertainty, and are relatively insensitive to risk information as it is frequently presented using percentages and frequency statements [259-263]. This issue is amplified when conditional probabilities are used. Many respondents process this information by creating risk clusters which only capture the gist of a risk [264-268], while some may ignore risk attributes entirely if the cognitive burden is too high (attribute non-attendance) [269]. If respondents misinterpret (or ignore) risk information then their preferences (as captured by the DCE) may not reflect the true value that they place upon the intervention in question, invalidating WTP estimates generated using these preferences. This could ultimately lead to suboptimal decision-making: interventions may be adopted when population WTP is less than total costs, or not adopted when the reverse is true.

Several studies have recently been published which suggest that respondents may better comprehend risk information in DCEs if it is presented using visual aids such as icon arrays (displays that use a matrix of icons to represent a risk) [263,264,270]. However, such approaches have not been widely applied to date [256,263,271] and empirical evidence to support a finding of improved risk comprehension is mixed [272]. If visual approaches do enhance the sensitivity of respondents to the magnitude of a risk, this may lead to greater

differences in risk perceptions, better capturing the true value of an intervention and potentially improving decision-making [270]. This effect could be captured in several ways, from changes in coefficient weights to larger standard errors around these estimates, reflecting the fact that a wider range of risk perceptions are now captured in the analysis. In addition, improved risk comprehension may directly motivate health-seeking behaviour in respondents, as causal linkages between behaviour and healthy outcomes are made more obvious [270]. A further benefit could be that estimates of WTP are generated which are more consistent with economic theory, an effect which has been noted in studies of risk presentation in the contingent valuation literature [263].

Studies comparing the use of visual approaches with frequency or probability statements in a DCE would therefore be a valuable addition to the limited literature in this area [263]. However, few have been undertaken to date. An exception is a study which investigated the impact of icon arrays on consumers' valuations of risk reductions, finding that these arrays did not reduce task complexity or error variance [27]. A literature review in the same paper identified no other studies conducting a similar test of risk presentation approaches, which provides further justification for conducting a comparative study. The use of genetic and genomic testing to guide treatment decisions in CLL represents an ideal case study for such an investigation, as patients who are asked to choose between different tests will have to evaluate complex risk information on test effectiveness and reliability in order to make their choices. A comparison of visual and non-visual approaches to presenting risk information to patients was therefore nested within the DCE presented in this chapter.

5.3 Designing the discrete choice experiment

5.3.1 Objectives

The primary objective of the DCE was to generate information on WTP for genetic and genomic testing. These estimates would then be used within a CBA to test the hypothesis that CLL patients value the information provided by genomic testing more highly than that provided by genetic testing. Although this 'bottom-line' estimate of WTP was the key output from the DCE, it was also important to understand differences in the WTP estimates for different attributes and levels, as this provides decision-makers with an indication of which test characteristics are valued by patients. Another consideration was whether patient preferences varied by socio-demographic characteristics, and whether the estimation of different choice models would impact on these results.

The secondary objective of the DCE was to evaluate the impact of presenting risk information in visual formats on both WTP and respondent understanding. A further consideration was whether any observed impacts were real effects (related to heterogeneous patient preferences) or were due to random differences in the error components of the choice models that were estimated.

5.3.2 The design process

Designing and conducting a DCE requires several tasks to be undertaken sequentially in order to generate a survey that can effectively address the stated research questions. The following section describes the steps taken during the development of this DCE. Prior to beginning this process, an exploratory literature review was undertaken to identify DCEs conducted in genetics or leukaemia that could inform the design of this DCE. A summary

of this review is provided in **Appendix 27**. The references that were identified provided a variety of information including examples of approaches to presenting risk information, suggestions of attributes and levels that might be important in this context and information on analytical approaches that are appropriate for DCEs in genomics.

Selecting a sampling population

The sampling population selected for this DCE was UK CLL patients as it is recommended that DCEs are undertaken in populations with experience in the area of interest [273]. Two populations were targeted to ensure that the preferences elicited by the DCE reflected the views of CLL patients with a range of characteristics (e.g. different locations and healthcare experiences), and to maximise the sampling population (given that CLL patients are typically elderly and thus response rates may be low). Population One included CLL patients attending outpatient clinics in the Oxford University Hospitals (OUH) NHS Trust (n=140), which provides acute care to a population of 650,000 people. Population Two included all patient members of the UK CLL Support Association (CLLSA, <http://www.clisupport.org.uk/>; n=982).

Establishing attributes and levels

Attributes and levels for the DCE were developed using several approaches. First, the references identified in the literature review in **Appendix 27** were reviewed to identify attributes used in previously published DCEs in genetics or leukaemia. Three haematologists at the OUH NHS Trust reviewed these attributes, identifying nine that were potentially relevant: those listed in **Table 5.1**, plus 'Number of blood samples required', 'Accuracy of the diagnosis', and 'Testing location'.

Interviews were then conducted with 15 randomly selected CLL patients at the OUH NHS Trust, who were asked to rate how important each attribute would be if they were deciding whether to undergo pre-treatment testing. The six highest ranked attributes were taken forward, based on how many would be manageable by this population, how many would be required to enable informed choices, and which would best capture the characteristics of current and future testing practice. These attributes are listed in **Table 5.1**. Further details of these interviews are provided in **Appendix 28** and **Appendix 29**.

Four clinically-feasible levels were identified for each attribute (informed by the patient and clinician interviews, literature searches and test characteristics). Potential interactions were identified between the ability of the test to predict who will not respond to the usual chemotherapy treatment (EFFECT) and COST, and also test reliability (REL) and COST (more expensive tests may be perceived to be better quality). Time to receive the test result (TIME), COST, EFFECT and REL were assumed to be linear (i.e. marginal utility is constant across the different levels that an attribute can take) and coded as continuous variables. Length of time clinicians spend describing the test (INFO) and type of clinician who explains the test result (WHO) were effects coded (the levels may be proxies for quantity or quality of information provided, so the effect may not be linear). **Table 5.1** also describes the expected impact on patient utility of an increase in the level of each attribute.

Table 5.1: Discrete choice experiment attributes and levels

Attribute	Level 1	Level 2	Level 3	Level 4	Expected impact on utility of an increase in the level of the attribute
Time to receive the test result [TIME]	5 days	8 days	11 days	14 days	Negative
Cost of the test [COST]	£130	£260	£400	£600	Negative
Ability of the test to predict who will not respond to the usual chemotherapy treatment [EFFECT]	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment	Positive
Test reliability [REL]	<u>2</u> out of every <u>100</u> tests provide an incorrect result	<u>4</u> out of every <u>100</u> tests provide an incorrect result	<u>6</u> out of every <u>100</u> tests provide an incorrect result	<u>8</u> out of every <u>100</u> tests provide an incorrect result	Negative
Length of time clinicians spend describing the test to you ^a	5 minutes [INFO0]	10 minutes [INFO1]	15 minutes [INFO2]	20 minutes [<i>base level</i>]	Positive ^b
Type of clinician who explains the test result to you ^a	General practitioner [WHO0]	Specialist nurse [WHO1]	Junior hospital doctor [WHO2]	Consultant hospital doctor [<i>base level</i>]	Positive? ^c

^a Effects-coded variable. ^b Coded as categorical variable, but as levels were ordered by increasing time, a positive impact was predicted. ^c The levels for this attribute were deliberately ordered so that perceived knowledge about CLL increased from general practitioner to specialist nurse to junior hospital doctor to consultant hospital doctor. This may not necessarily translate into a positive increasing effect on utility, as other factors may also be important to patients.

Risk presentation formats

The graphical approach to be used for the presentation of risk information in the DCE was selected in a similar way to the attributes and levels. The most common format in the literature was a grid of one hundred icons, with shading used to illustrate risk information. This format was said to enable better processing of quantitative information and also minimise denominator neglect (where respondents focus on event incidence and ignore the number of times that the event could have happened) [261,266,274]. A recent systematic review supports this conclusion [263]. Within this basic framework a wide variety of icons were used. The five most commonly used icons and arrangements were taken forward to the patient interviews described in the previous section. In these interviews, patients were presented with ten examples of risk graphics (using the five formats mentioned previously, highlighted sequentially or non-sequentially) and asked to select the frequency statement that matched this graphic from four options. The format that returned the most correct responses – sequential displays using a grid of faces – was taken forward, on the basis that this enabled better processing of quantitative information by patients. Full details of these patient interviews are provided in **Appendix 28** and **Appendix 29**.

Two of the attributes selected for the final DCE were risk-based (ability of the test to predict who will not respond to the usual chemotherapy treatment [EFFECT] and test reliability [REL]), so risk graphics were used for both. **Figure 5.1** presents a sample question from the DCE, both with and without risk graphics.

Figure 5.1: Example discrete choice experiment question both with and without risk graphics

Consider the following features describing two tests (Test A or Test B). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	5 days
Cost of the test <u>to you</u>	£260	£600
Ability of the test to predict who will <u>not</u> respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result	<u>4</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	10 minutes	5 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	Junior hospital doctor

Which test would you prefer?

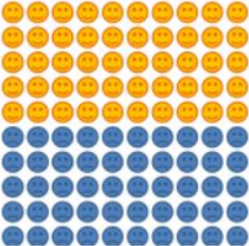
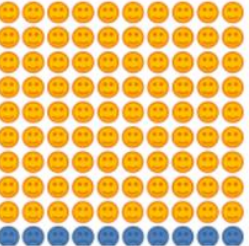
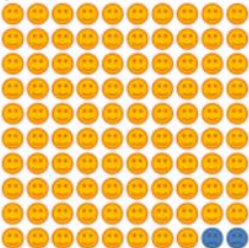
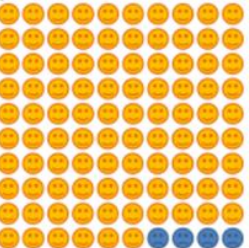
☐

TEST A

☒

TEST B

Consider the following features describing two tests (Test A or Test B). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	5 days
Cost of the test <u>to you</u>	£260	£600
Ability of the test to predict who will <u>not</u> respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result 	<u>4</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	10 minutes	5 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	Junior hospital doctor

Which test would you prefer?

☐

TEST A

☒

TEST B

Experimental design

The chosen attributes and levels were used to design a DCE in which respondents were presented with choice sets containing two alternatives (Test A vs. Test B). An alternative design including an opt-out was considered to evaluate potential test uptake. To determine whether this would add additional complexity for limited benefit, a pilot DCE was generated with a two-stage design (Test A vs. Test B, then Chosen test vs. No test). This was completed by a convenience sample of 14 members of the University of Oxford, and six members of the general public. 'No test' was selected in only 3.9% of choices. Given that genetic testing is currently recommended for UK CLL patients [158] (hence their ability to choose 'no test' is limited), this was judged to be sufficiently low that the additional complexity associated with adapting the DCE design and analysis outweighed the additional information that was likely to be provided concerning test uptake.

Respondents were asked to complete 16 choice sets. This decision was informed by studies which suggest that the number of choice sets does not impact on response rates [273,275]. The two alternatives in each choice set were unlabelled to ensure that respondents based decisions on attribute levels, not test names. At no point were tests called 'genetic' or 'genomic'; they were only defined by attributes and levels.

A second pilot was then undertaken to generate priors to inform the main design. A fractional factorial design was produced using Ngene [276]. Twelve CLL patients at the OUH NHS Trust completed this pilot. Multinomial logit (MNL) regression (sometimes called conditional logit) was used to analyse the choice data in Stata [277]. Coefficient signs were mostly as expected, with counter-intuitive results generally not significant. These priors

were incorporated into the final experimental design. To permit the investigation of preference heterogeneity, this design used a model averaging approach which incorporated more flexible model specifications (e.g. mixed logit) [278]. Two DCE surveys were produced using the same design, with and without risk graphics. Additional information on the pilot DCE and the final experimental design is provided in **Appendix 30**.

Constructing the survey

The final survey comprised four sections. Section one provided respondents with background information on genetic and genomic testing and described the attributes and levels. In section two, respondents ranked attributes in order of preference. Section three contained the DCE task, preceded by a 'rationality check' choice in which Test A contained the worst levels for each attribute and Test B contained the best. Section four collected information on respondent characteristics and clinical details. Respondents were also asked for their opinions on genetic testing, to report their health status using the EuroQol five dimensions survey instrument, and to rate the difficulty of the survey.

Administering the survey

The survey was administered in two forms. Patients at the OUH NHS Trust (n=140) were asked to complete a paper survey during an outpatient appointment. Surveys were distributed sequentially (patient one received a survey with risk information presented using graphics, patient two received a survey without graphics etc.). CLLSA patient members received either paper (n=148) or electronic (n=834) versions of the survey. Again, questionnaires were distributed sequentially, after sorting members into alphabetical order. All those wishing to return a paper version of the survey received a prepaid

envelope. Patients completing the DCE electronically returned their survey by email. All patients received two reminders. Apart from the planned differences in risk presentation and minor differences in the instructions for completion, all surveys were identical across the different patient groups and formats. Versions of the DCE survey with and without risk graphics can be found in **Appendix 31** and **Appendix 32**. Data collection took place from July to October 2013. Ethical approval for the survey was sought from the National Research Ethics Service and the OUH NHS Trust R&D office. Both bodies stated that ethical approval was not required.

Data analysis

Data analysis was undertaken using Stata [277]. In Model A, the choice data for all respondents was modelled using MNL regression, testing main effects only. This initial model assumes that preferences are homogenous across individuals. Differences by study design (including respondent group and completion method) were evaluated, and models were estimated separately for respondents who were exposed to different methods of risk presentation. To test model validity, models were estimated with and without respondents who failed the rationality check. Choices were also evaluated separately for respondents who stated that the survey was difficult. Missing data and the existence of dominant preferences were quantified in the same manner. In addition, the presence of dominant preferences was explored by examining if respondents always chose the test with the best level for a specific attribute [256]. All tests were conducted separately for respondents who were exposed to visual and non-visual methods of risk presentation.

Model A was then adapted to test the significance of interactions between attributes and respondent characteristics. The characteristics that were examined included gender, age, occupation, children, income, time since CLL diagnosis, experience of chemotherapy treatment and genetic testing, whether the respondent had favourable opinions about genetic testing, and health status. Subsequent analyses accounted for the limitations of MNL models [279] by fitting alternative functional forms (including mixed logit and latent class models), which allowed for preference heterogeneity. Different specifications were compared using the Akaike, Bayesian and Consistent information criterion to identify the most appropriate specification (hereafter Model B). This enabled an investigation into whether the use of different choice models modified the impact of visual risk information.

For all models the marginal rate of substitution (the ratio between cost and other attributes) was calculated, WTP for testing was estimated, and ratios between attributes were calculated to understand how much of one attribute respondents were willing to give up to get more of a second attribute. Confidence intervals around WTP estimates were calculated using the Stata module 'wtp'. Finally, it was noted that the models for visual and non-visual methods of risk presentation may produce different coefficient estimates, not necessarily because preferences differ across the two populations, but because of differences in the variance of their error components. A scale parameter test was therefore conducted for Models A and B to isolate the true driver of any differences and to establish whether there is parameter equality across the different versions of the DCE [280].

5.4 Discrete choice experiment results

5.4.1 Survey metrics and respondent characteristics

Eighty paper surveys (response rate (RR): 54%) and 70 electronic surveys (RR: 8%; overall RR: 15%) were received from CLLSA members. Of the 140 patients asked to complete a survey at the OUH NHS Trust, 101 were recruited (39 already participated in study/unwilling), with 69 paper surveys returned (RR: 68%). The overall response rate for paper surveys was 60%. 219 CLL patients completed a DCE in total (RR: 20%). Of these, 111 (51%) completed a survey with risk graphics. Response rates were similar with and without graphics (21% and 20%, respectively), for both paper and electronic surveys.

The age, gender distribution and employment status of respondents were typical of CLL patients (**Table 5.2**) [158]. On average, respondents were diagnosed with CLL six years before completing the survey, with half undergoing at least one course of chemotherapy. Of these, 44% had at least one inpatient stay due to chemotherapy treatment, with 12% currently undergoing chemotherapy. One in five reported undergoing genetic testing, with 96% in favour of pre-treatment genetic or genomic tests (**Appendix 33**). The current quality of life of respondents matched UK population norms for this age group [281,282].

Most respondents (97%) passed the rationality check. This result was unaffected by use of risk graphics. Few respondents (9%) rated the DCE as difficult to complete (10% with risk graphics, 7% without). Prior to undertaking the DCE, the most important attribute to respondents was EFFECT and the least important was COST (**Appendix 33**). Only 0.8% of choice questions were not answered (0.3% with risk graphics, 1.3% without). Missing questions were mostly consecutive (i.e. respondents likely turned over two pages at once),

so were assumed to be missing at random. Finally, 96 respondents (44%) made dominant choices (50% with risk graphics, 37% without). However, 88 (92%) were dominant for EFFECT, which exhibited overlap in 25% of choice sets (hence it was easier to be dominant on this attribute).

Table 5.2: Discrete choice experiment respondent characteristics

Variable	Category	Value	N
Male	-	117 (53%)	219
Mean age in years (SD)	-	65.7 (10.2)	205
Employment status	Employed Retired Other	67 (31%) 138 (63%) 14 (6%)	219
Mean age on leaving full-time education in years (SD)	-	18.7 (4.5)	205
Mean household income per annum (£)	-	30,646	201
Mean number of people per household (SD)	-	2.0 (0.9)	212
Mean number of children (SD)	At home Away from home	0.4 (0.8) 1.7 (1.3)	219
Mean travel time to hospital (minutes)	-	35.5	217
Mode of transport used to reach hospital	Hospital arranged Public transport Private transport	3 (1%) 34 (16%) 178 (83%)	215
Mean time since diagnosis in years (SD)	-	6.0 (4.5)	215
Number of respondents who have undergone chemotherapy at least once	-	98 (45%)	216
Average number of courses of chemotherapy in these respondents (SD)	-	1.3 (0.8)	93
Number of respondents who have had an inpatient stay after chemotherapy treatment	-	42 (44%)	96
Number of respondents undergoing chemotherapy currently	-	11 (12%)	94
Number of respondents who had previously undergone genetic testing	-	44 (20%)	216
Mean EQ-5D index score (SD)	-	0.823 (0.225)	217
Mean EQ-VAS score (SD)	-	74.2 (17.2)	210

EQ-5D = EuroQol five dimensions survey instrument; SD = standard deviation; VAS = visual analogue scale.

5.4.2 Model A results

Table 5.3 presents the results for Model A, which tested main effects and differences by study design. Initial analyses considered the impact on model performance of excluding patients who failed the rationality check, made dominant choices, described the DCE as difficult, or missed choice tasks. The only variation that increased pseudo-R², changed at least one coefficient from insignificant to significant and retained sufficient choice data was excluding those who described the DCE as difficult (n=19). These patients – who were older (71.7 years), left full-time education earlier (17.7 years), and missed more questions (2.0%) than the average respondent – were therefore excluded from the analysis.

The TIME, COST, EFFECT (test effectiveness) and REL (test reliability) coefficients all had the expected sign: respondents prefer tests that are more effective, more reliable, cheaper, and return results quickly. The signs for the INFO coefficients indicated that only a 15 minute appointment has a positive impact on utility. The signs for the WHO coefficients indicated a utility gain only when results are explained by specialist nurses or consultant hospital doctors. Improvements in test effectiveness are valued, with patients willing to pay £24 for a 1% increase in the proportion of chemotherapy non-responders identified. However, process attributes are also valued: patients are willing to pay £27 to reduce time to receive test results by one day. The coefficient ratios indicated that respondents would be willing to wait an extra 29 days for test results if an additional one-third of chemotherapy non-responders could be identified. Alternatively, respondents would tolerate a genomic test being wrong 8% of the time to receive this information. Finally, receiving results from a consultant hospital doctor instead of a general practitioner is equivalent to a 15% increase in test effectiveness.

Table 5.3: Regression results and willingness-to-pay estimates

Attribute ^a	All respondents		Risk graphics		No risk graphics	
	Coefficients	WTP (95% CI)	Coefficients	WTP (95% CI)	Coefficients	WTP (95% CI)
<i>Model A (MNL)</i>						
TIME	-0.0461**	-£27.19 (-£38.67 to -£15.71)	-0.0486**	-£30.72 (-£50.42 to -£11.01)	-0.0452**	-£25.87 (-£40.64 to -£11.10)
COST	-0.0017**	N/A	-0.0016**	N/A	-0.0017**	N/A
EFFECT	0.0404**	£23.79 (£16.06 to £31.51)	0.0468**	£29.59 (£14.48 to £44.70)	0.0354**	£20.25 (£11.47 to £29.03)
REL	-0.1584**	-£93.38 (-£125.79 to -£60.98)	-0.1251**	-£79.07 (-£125.06 to -£33.07)	-0.1844**	-£105.63 (-£153.12 to -£58.14)
INFO0	-0.1367*	-£80.59 (-£165.73 to £4.54)	-0.1927	-£121.79 (-£270.88 to -£27.31)	-0.0838	-£48.03 (-£152.78 to £56.73)
INFO1	-0.0319	-£18.83 (-£81.66 to £43.99)	-0.0731	-£46.20 (-£155.98 to £63.59)	-0.0088	-£5.07 (-£84.81 to £74.67)
INFO2	0.2624**	£154.69 (£55.61 to £253.76)	0.3863**	£244.14 (£54.58 to £433.69)	0.1689	£96.75 (-£18.99 to £212.49)
INFO3 ^b	-0.0937**	-£55.26	-0.1205**	-£76.15	-0.0762**	-£43.66
WHO0	-0.8399**	-£495.10 (-£646.56 to -£343.64)	-0.8203**	-£518.38 (-£772.24 to -£264.53)	-0.8652**	-£495.58 (-£695.04 to -£296.12)
WHO1	0.5071**	£298.92 (£181.14 to £416.71)	0.5235**	£330.84 (£130.60 to £531.08)	0.4932**	£282.50 (£131.93 to £433.06)
WHO2	-0.2588**	-£152.54 (-£251.26 to -£53.81)	-0.3362**	-£212.44 (-£395.14 to -£29.74)	-0.1909*	-£109.33 (-£225.45 to £6.79)
WHO3 ^b	0.5915**	£348.71	0.6329**	£399.98	0.5629**	£322.41
Pseudo-R ²	0.4101		0.4677		0.3688	
<i>Model B (MXL)</i>						
TIME	-0.0556**	-£31.05 (-£47.13 to -£14.97)	-0.0720**	-£41.37 (-£71.28 to -£11.45)	-0.0537**	-£29.11 (-£49.04 to -£9.18)
COST	-0.0018**	N/A	-0.0017**	N/A	-0.0018**	N/A
EFFECT	0.0963**	£53.79 (£32.05 to £75.54)	0.1047**	£60.15 (£22.82 to £97.48)	0.0807**	£43.78 (£20.06 to £67.50)
REL	-0.2789**	-£155.73 (-£220.09 to -£91.38)	-0.1826**	-£104.91 (-£174.70 to -£35.11)	-0.2920**	-£158.44 (-£244.13 to -£72.75)
INFO0	-0.1086	-£60.66 (-£160.61 to £39.30)	-0.1777	-£102.06 (-£273.27 to £69.14)	-0.0602	-£32.67 (-£160.73 to £95.39)
INFO1	-0.0531	-£29.67 (-£117.92 to £58.58)	-0.1591	-£91.39 (-£249.31 to £66.52)	0.0195	£10.60 (-£103.65 to £124.84)
INFO2	0.1510	£84.31 (-£30.07 to £198.69)	0.4637**	£266.35 (£21.45 to £511.26)	0.0270	£14.66 (-£119.60 to £148.92)
INFO3 ^b	0.0108**	£6.02	-0.1269**	-£72.90	0.0137**	£7.41
WHO0	-0.9739**	-£543.81 (-£753.89 to -£333.74)	-1.0362**	-£595.25 (-£956.63 to -£233.86)	-1.0633**	-£576.92 (-£868.80 to -£285.05)
WHO1	0.4756**	£265.58 (£127.83 to £403.32)	0.5325**	£305.91 (£73.80 to £538.02)	0.4753**	£257.87 (£75.51 to £440.23)
WHO2	-0.3747**	-£209.22 (-£354.08 to -£64.36)	-0.4098**	-£235.39 (-£480.88 to £10.11)	-0.3315**	-£179.86 (-£361.97 to £2.26)
WHO3 ^b	0.8730**	£487.46	0.9135**	£524.72	0.9195**	£498.91

95% CI = 95% confidence interval; MNL = multinomial logit; MXL = mixed logit; WTP = willingness-to-pay. * P-value < 0.05; ** P-value < 0.01. TIME = time to receive test result; COST = cost of the test; EFFECT = ability of the test to predict who will not respond to the usual chemotherapy treatment; REL = test reliability; INFO = length of time clinicians spend describing the test to you (0=5min, 1=10min, 2=15min, 3=20min); WHO = type of clinician who explains the test result to you (0=general practitioner, 1= specialist nurse, 2=junior hospital doctor, 3=consultant hospital doctor). ^a The coefficient for the alternative specific constant was not significant in the initial regressions (indicating no respondent left-right bias) so was dropped from subsequent regressions. ^b WTP 95% CIs cannot be calculated for the base levels of the effects coded variables.

When Model A was estimated separately for different survey formats, pseudo- R^2 was higher (0.5362 vs. 0.3335) for the model containing DCEs completed electronically (although the signs and significance of most coefficients did not change). WTP for a 1% increase in test effectiveness was higher when the DCE was completed electronically (£24.33 vs. £20.92), as was WTP for a 1% improvement in test reliability (£93.33 vs. £82.99). When Model A was estimated separately by sample population, pseudo- R^2 was higher for the model for CLLSA respondents (0.4510 vs. 0.2735), although the signs and significance of most coefficients again remained the same. WTP for a 1% increase in test effectiveness was higher for CLLSA respondents (£32.21 vs. £14.07), while WTP for a 1% improvement in test reliability was higher for respondents from the OUH NHS Trust (£121.06 vs. £58.53).

Incorporating attribute interactions

Appendix 34 presents the results of an extension to Model A which evaluated attribute interactions. Neither of the planned interactions (COST x EFFECT, COST x REL) were significant, so were excluded from the final specification. However, when significant interactions between attributes and respondent characteristics were added, pseudo- R^2 increased and the significance of INFO0 improved. Notable interactions included MALE x INFO0 (men have a stronger preference for shorter appointments), MALE x WHO0 (men have a stronger preference for receiving results from general practitioners), and INCOME x REL (higher income patients have a weaker preference for more reliable tests).

Adapting Model A to consider the impact of different methods of risk presentation

To evaluate the impact of presenting risk information visually, the MNL model for all respondents was first extended to include interactions between the use of risk graphics

and the two risk attributes (EFFECT and REL). Both interactions were significant, and pseudo- R^2 increased from 0.4101 to 0.4167. A MNL model was then estimated separately for respondents who completed DCEs with and without risk graphics. These results are reported in **Table 5.3**. The signs and significance of most coefficients remained the same as in the model for the entire sample, and both models fitted the choice data well. The key coefficients are those for the attributes where the presentation of risk information varied between surveys. The EFFECT coefficients indicated that an increase in test effectiveness has a greater positive effect on utility when risk graphics are used. The standard error for this coefficient was also higher when risk graphics were used (0.0023 vs. 0.0019; 0.0014 in the model for the entire sample). Conversely, the REL coefficients indicated that a decline in test reliability, such as a change from 4% of tests producing an incorrect result to 8%, has a smaller negative effect on utility when respondents see risk information presented graphically. The standard error for this coefficient was, however, again higher when risk graphics were used (0.0229 vs. 0.0198; 0.0148 in the model for the entire sample). This effect is reflected in the WTP estimates. WTP for a 1% increase in test effectiveness was higher when risk graphics were used (£29.59 vs. £20.25), while WTP for a 1% improvement in test reliability was higher when risk graphics were not used (£105.63 vs. £79.07).

Several of the coefficients and WTP estimates for the non-risk attributes were also modified by the use of risk graphics. Respondents were willing to pay £399.98 for results to be explained by a consultant hospital doctor when risk graphics were used, but £322.41 when non-visual methods were used. Conversely, WTP for a junior hospital doctor fell from -£109.33 to -£212.44 when risk graphics were used. Preferences for different appointment lengths also varied: WTP for a 15 minute appointment was £244.14 when risk graphics

were used, falling to £96.75 without graphics (although this latter coefficient was not significant). The only coefficient that remained relatively stable regardless of how risk information was presented was that for the cost attribute. Finally, the Swait and Louviere scale parameter test rejected the hypothesis of parameter equality [280], suggesting that there is preference heterogeneity: respondents completing the choice tasks with risk graphics have different preferences to those completing the tasks without risk graphics.

5.4.3 Model B results

Table 5.4 presents a comparison of alternative specifications for Model A. These results suggest that alternative specifications provide a better fit than MNL for Model A, with the most appropriate specification being either a mixed logit or two-class latent class approach. As limited information is available to characterise the latent classes in this main effects model, the mixed logit approach (hereafter Model B) is taken forward. A comparison of alternative specifications was also undertaken for the version of Model A which incorporated attribute interactions. These results are presented in **Appendix 34**.

Table 5.4: Alternative specifications for Model A

Specification	AIC	BIC	CAIC
Multinomial logit	2,553	2,620	N/A
Alternative-specific conditional logit	2,555	2,629	N/A
Mixed logit (hereafter <u>Model B</u>)	2,134	2,268	N/A
Latent class (two classes)	2,184	2,184	2,316
Generalised multinomial logit	2,137	2,286	N/A

AIC = Akaike information criterion; BIC = Bayesian information criterion; CAIC = Consistent information criterion.

The results for Model B are presented in **Table 5.3**. In the model for all respondents, the significance of most coefficients remains the same as in Model A, although three are no

longer significant: INFO0, INFO1 and INFO2. Notable changes in WTP include EFFECT (which increases from £23.79 to £53.79) and REL (which decreases from -£93.38 to -£155.73). Overall, this model correctly predicted 81 % of respondent choices.

Adapting Model B to consider the impact of different methods of risk presentation

Model B was also adapted to consider the impact of risk graphics. These results are presented in **Table 5.3**. The signs and significance of most coefficients remained the same as in the model for the entire sample, with three exceptions. First, a ten minute appointment (INFO1) has a negative impact on utility in the model for the entire sample and the model using risk graphics, but not in the model without graphics. Second, a 20 minute appointment (INFO3) has a positive impact on utility in the main model and when non-visual methods are used, but not when risk graphics are used. Finally, the coefficient for a 15 minute appointment (INFO2) is only significant when risk graphics are used.

As in Model A, the EFFECT coefficients indicated that an increase in test effectiveness has a greater positive effect on utility when risk graphics are used. In addition, the standard error for this coefficient was higher when risk graphics were used (0.0108 vs. 0.0093; 0.0081 in the model for the entire sample). The behaviour of the REL coefficients was also the same in both models: a decline in test reliability has a smaller negative effect on utility when respondents see risk information presented graphically. The standard errors for this coefficient were 0.0367 (risk graphics), 0.0425 (no risk graphics) and 0.0311 (entire sample). WTP for a 1% increase in test effectiveness was higher when risk graphics were used (£60.15 vs. £43.78), while WTP for a 1% improvement in test reliability was higher

when risk graphics were not used (£158.44 vs. £104.91). Regardless of risk presentation format, WTP for improvements in EFFECT and REL was higher in Model B than Model A.

Again, the coefficients and WTP estimates for most of the non-risk attributes (except COST) were modified by the use of risk graphics. As per Model A, this effect was observed for WTP for a 15 minute appointment: this was £266.35 with risk graphics, but only £14.66 without. WTP for a junior hospital doctor also fell from -£179.86 to -£235.39 when risk graphics were used. WTP values for the non-risk attributes were also generally higher in Model B than in Model A. Finally, the scale parameter test again rejected the hypothesis of parameter equality, indicating that respondents completing the choice tasks with risk graphics have different preferences to those completing the tasks without risk graphics.

5.4.4 Impact of risk graphics on willingness-to-pay for genomic testing

The results of Models A and B with respect to the use of risk graphics were broadly consistent in terms of the direction of respondents' preferences. However, the strength of these preferences varied. When risk graphics were used, if Model A returned a positive coefficient, Model B always returned a larger positive coefficient. For negative coefficients, Model B generally returned a larger negative coefficient than Model A i.e. a smaller absolute number. This relationship was less consistent when risk graphics were not used. The same pattern of results was observed when the WTP estimates generated by the two models were compared with respect to the use of risk graphics. The overall impact of variations in WTP by use of risk graphics can be evaluated by considering WTP for a hypothetical specification of genomic testing in CLL. **Table 5.5** presents WTP estimates for both choice models, by use of risk graphics, for the following specification of testing:

- Time to receive test result: 14 days
- Ability of the test to predict who will not respond to the usual chemotherapy treatment: the test can identify 67 out of every 100 patients (approximately two-thirds) who will not respond to usual treatment
- Test reliability: 2 out of every 100 tests provide an incorrect result
- Length of time clinicians spend describing the test: 15 minutes
- Type of clinician who explains the test result: Consultant hospital doctor

Table 5.5: Willingness-to-pay for a hypothetical specification of genomic testing in chronic lymphocytic leukaemia

Model	Willingness-to-pay			Ratio of graphics to no graphics
	Whole sample	Risk graphics	No risk graphics	
Model A (MNL)	£1,529.91	£2,038.43	£1,202.47	1.7
Model B (MXL)	£3,429.54	£4,032.12	£2,722.41	1.5
Ratio of B to A	2.2	2.0	2.3	

MNL = multinomial logit, MXL = mixed logit.

In both models, WTP for a hypothetical specification of testing is 1.5-1.7 times higher when risk graphics are used compared to no graphics. There is also a difference in WTP between the two models: estimates are 2.0-2.3 times higher in Model B compared to Model A.

5.4.5 Discussion

Overview

The results of the DCE suggest that patients prefer tests that are more effective, more reliable, cheaper, and return results quickly, preferably in a 15 minute appointment with a clinician who is perceived to be a CLL expert. Both clinical and process-related attributes are important to patients, with test effectiveness and speed of result ranked highly. Patient

characteristics may modify these findings, with male patients (who account for two-thirds of CLL incidence and have worse outcomes [283]) having a stronger preference than women for shorter appointments with general practitioners (rather than with hospital-based clinicians). This supports the suggestion that health-seeking behaviour differs between men and women [284,285]. Overall, respondents expressed a preference for combinations of attributes and attribute levels that more closely reflect genomic testing than genetic testing. Indeed, respondents were willing to make notable trade-offs for the extra information provided by genomic testing (including waiting a month longer for results if genomic testing can identify an additional one-third of chemotherapy non-responders).

These findings match those of other DCEs that have evaluated genetic and genomic tests. Regier *et al.* evaluated the use of genomic testing to identify the genetic causes of developmental disability, reporting that both the diagnostic rate and waiting time for test results were important to families [258,286]. Herbild *et al.* evaluated WTP for pharmacogenetic testing in the treatment of depression and found that respondents placed a high value on how the intervention was delivered [287]. Finally, Najafzadeh *et al.* measured the preferences of patients and the general public for genomic testing to determine drug response, finding that both clinical (e.g. test effectiveness) and process-related attributes (e.g. genetic test procedure) are important to respondents [288].

Risk presentation

Importantly, the preferences of respondents who were presented with risk graphics differed from those of respondents who were presented with the same risk information without graphics. These differences fed through to estimates of WTP, suggesting that the

two methods of risk presentation measure different cognitive processes. The differences between the two presentational formats were manifested in several ways. In terms of test effectiveness, the use of graphics increased WTP for more effective tests and led to less precise coefficient and WTP estimates. In terms of test reliability, the use of graphics decreased WTP for more reliable tests, with an inconclusive impact on the precision of coefficient and WTP estimates. The differing impact of risk graphics could be explained by the way that attributes were framed in the DCE. For test effectiveness, a higher value (more patients identified who would not respond to chemotherapy) was a definitively 'good' outcome. For test reliability, a higher value (more tests provide an incorrect result) was a definitively 'bad' outcome. Previous studies have also noted this effect [265].

The coefficients for attributes without risk components (INFO and WHO) were also modified by the use of risk graphics. One potential explanation for this finding is that use of graphics fundamentally impacts on respondents' understanding of the risk information that is presented to them. A second potential explanation is that changes in preferences are driven by changes in the design of the DCE rather than an improved understanding of risk. When graphics are used, the risk attributes visually dominate the choice question, taking up more room within the choice set than non-risk attributes. Respondents may therefore pay more attention to risk attributes in this scenario. This may result in improved understanding, or respondents may instead simply view these attributes as being more important because they dominate the page, which feeds through to coefficient estimates. A third consideration is that the survey with risk graphics was longer than the survey without graphics, hence survey fatigue could be higher in respondents who view graphics.

However, previous studies have suggested that the provision of more choice sets (which may be a proxy for survey length) does not impact on the results of DCE surveys [273,275]. Some aspects of the DCE were unaffected by the use of graphics, including response rate and comprehension (as measured by the rationality check question, difficulty ratings and amount of missing data). However, respondents may be more likely to make dominant choices when risk graphics are used.

Only one previous study has investigated the consequences of presenting risk information in a DCE in visual and non-visual formats [27]. This study also reported that risk graphics did not impact on task complexity, and concluded that graphics help respondents to understand changes in risk. This study also noted that WTP was higher when risk graphics were used, and concluded that the mixed logit approach performed best.

Discrete choice experiment limitations

The DCE has several limitations. First, a broad cross-section of patients was targeted by using two sampling populations. However, some patients may be under-represented and the sample may therefore be characterised by selective non-response. The most severely ill may have been unable to participate, while the healthiest patients at the OUH NHS Trust (who only have annual check-ups) may not have been sampled. Furthermore, CLLSA members differ from non-members. However, given that respondent characteristics were typical of CLL patients [158], these factors are unlikely to have affected the DCE results.

Second, the DCE response rate was low, particularly for CLLSA members who completed electronic surveys. DCE response rates have been decreasing over time [289] and low

response rates are no longer unusual, particularly in elderly populations and for electronic DCE surveys [290]. Furthermore, evidence suggests that response rates are lower when more than four attributes are included in a DCE, and when there is no opt-out [289]. However, most published DCEs include between 100 and 300 respondents and the sample size of 219 respondents in this DCE does fall within this range [291,292]. It has been noted previously that response rates are positively related to the perceived benefits of completing a DCE survey [289] hence it is possible that the low response rate in this survey may have impacted on the DCE results. Specifically, if those who completed the survey were primarily patients who were more aware of the benefits of testing (e.g. patients who had previously undergone chemotherapy and were thus familiar with the consequences of treatment) this might have biased WTP estimates upwards. However, as 55% of respondents had not undergone chemotherapy and respondent quality of life matched population norms, the low response rate is unlikely to have impacted on the results in this manner.

Third, an opt-out design was tested in a pilot which used a convenience sample of non-CLL patients and which found that the opt-out option was rarely selected. Ideally, this pilot would have used CLL patients, but the small pool of patients at the OUH NHS Trust precluded this approach. As CLL patients are rarely able to choose 'no test' in the UK, and general population respondents may consider an opt-out to be a more viable choice than CLL patients, it is likely that a pilot of the opt-out design in CLL patients would have seen even fewer respondents selecting the opt-out option, providing further justification for this design decision.

Fourth, the TIME, COST, EFFECT and REL attributes were assumed to be linear and coded as continuous attributes to permit the use of the DCE results to characterise genomic testing practice (whose attributes were unknown at the time the DCE was designed) in future cost-benefit analyses. It is, however, possible that the relationship between these attributes and utility is not linear, hence this is a potential weakness of this study.

Fifth, almost all respondents were in favour of using pre-treatment genetic or genomic testing in CLL; this could have impacted on the WTP estimates.

Sixth, prior to undertaking the survey respondents were asked to rate the importance of each attribute. However, similar information was not collected at the end of the survey, which would have permitted further assessment of the impact of risk graphics.

Seventh, 19 respondents described the DCE as difficult and were excluded from the analysis, which improved model fit. These respondents were excluded because it was not clear that they understood the choice questions. These patients also missed out a greater proportion of the choice questions compared to those who said the DCE was not difficult, hence their responses provided limited additional information on trade-offs between test attributes. This decision was considered carefully as it has been suggested that the removal of such respondents is inappropriate [273]. As few respondents were affected, this decision is unlikely to have affected the main conclusions in this section. Furthermore, by removing respondents who may not have understood the choice questions, these conclusions may be more robust.

Eighth, these results may be specific to the graphical risk presentation approach used in this study, although these results do closely match those of Erdem *et al.*, who used different graphics [27]. Given that respondents have differing levels of understanding of risk formats (which will feed through to choice behaviour), a study which evaluates multiple methods of risk presentation could better demonstrate the impact of graphics.

Ninth, this DCE considered how differences in the presentation of risk information (and subsequently the interpretation of this information by respondents) might impact on WTP. However, this is one specific variation in attribute presentation and it is possible that alternative variations might have also affected the interpretation of attribute information by respondents, changing WTP estimates in a similar manner. For example, preferences for the type of health care professional who explains test results to patients might have varied had additional information been provided on the characteristics of these clinicians (e.g. average age, length of training). Without additional information on the magnitude of changes in WTP as a consequence of alternative presentational variations, it is not possible to fully place the observed differences in WTP by use of risk graphics into context.

Tenth, the selection of attributes for the DCE was determined by a literature review, followed by consultations with clinicians and then patients. It is possible that the ordering of this step in the design process may have biased the DCE design towards the use of attributes which are more important to clinicians than to patients. Evidence has recently emerged that clinicians and patients may have different preferences regarding pharmacogenetic testing, with clinicians potentially valuing process-related attributes more highly than patients [256,293]. The final DCE design might therefore have been

different had patients been consulted prior to clinicians. This limitation may be mitigated somewhat by the fact that the attributes presented to clinicians were informed by a review of previously published DCEs which were primarily conducted in patients. However, as a result of this design decision the CBA presented in sections 5.5 and 5.6 is quite narrowly specified. This complicates the comparison between the results of the welfarist and extra-welfarist economic evaluations in Chapter Six, although the 'what if' benefit transfer analysis presented in section 6.2.4 concludes that this design decision is unlikely to have changed the overall results of the study.

Eleventh, the results of DCEs which use a cost attribute in the context of healthcare decision-making in the UK NHS should be interpreted cautiously as healthcare is almost always free at the point of use in the UK system. The description of the cost attribute in the DCE survey asked respondents to imagine that they lived in a country in which they had to cover the cost of medical tests out of their own pocket. As there were no protest responses to the DCE, and only three respondents exhibited dominant preferences for the cost attribute, it appears unlikely that this design decision impacted on the DCE results.

Finally, this result reflects a specific clinical context. However, these findings are likely to be more generalisable than this, particularly for other chronic cancers and chronic diseases outside of cancer with similarities in clinical practice (e.g. timing of treatment decisions).

5.5 Designing the cost-benefit analysis

Sections 5.2-5.4 have described the design, analysis and results of a DCE which estimated the WTP of CLL patients for the use of genetic and genomic testing to guide chemotherapy

treatment decisions. As the overall aim of this chapter is to consider the costs and consequences of these tests from a welfarist perspective, the following two sections present a CBA which used the DCE results to achieve this objective. This section details the design of this CBA. The basic analytical framework has been described in Sections 2.2.2 and 5.2.1, so this section focuses on three key steps in the design process: monetising the benefits of genetic and genomic testing, estimating costs for use in the CBA, then bringing costs and benefits together and determining which test has the highest net benefit.

5.5.1 Monetising the benefits of genetic and genomic testing

The DCE results were used to monetise the benefits of genetic and genomic testing to conduct this CBA. This involved selecting the comparators to be considered in the analysis, specifying attribute levels for each comparator, then valuing these attribute levels.

Selecting the comparators

Although the CEA and CUA presented in Chapter Four considered five comparators, reflecting different configurations of genetic and genomic testing followed by treatment, the DCE was more narrowly specified, only asking patients to make choices between different types of genetic or genomic test. This approach was taken as it was felt that patients would find it difficult to express preferences for both testing and treatment in the same DCE, and the DCE would be prohibitively complex if all relevant testing and treatment attributes were included. The limited literature on the use of DCEs to conduct CBAs in healthcare settings suggests that any WTP estimates that are generated by the DCE should be driven entirely by how the cost attribute is framed [86]. As this attribute only reflects

the cost of testing (not treatment following testing), the CBA consequently only considers the net benefits associated with two comparators: genetic testing and genomic testing.

Specifying the attribute levels for each comparator

To allow WTP for genetic and genomic testing in this context to be calculated, both forms of testing must be specified in terms of the attributes and levels in the DCE. Clinicians in the OUH NHS Trust were consulted in order to specify the attribute levels for each test, and these are presented in **Table 5.6**.

Table 5.6: Attribute levels for genetic and genomic testing in the cost-benefit analysis

Attribute	Current genetic testing practice	Future genomic testing practice
Time to receive test result	14 days	14 days
% patients identified as FCR non-responders	6.77%	16.67%
% times FCR non-responders are identified incorrectly	22.22%	18.18%
Length of appointment	15 minutes	15 minutes
Type of clinician	Consultant hospital doctor	Consultant hospital doctor

FCR = rituximab, cyclophosphamide and fludarabine.

The same levels were used for length of appointment and type of clinician for both tests as these attributes are unlikely to change (a sensitivity analysis did consider the possibility that a longer appointment would be required to present genomic testing results). Similarly, time to receive test result also remained the same in both cases. This is a potentially conservative assumption as genomic test results will likely be returned more quickly in the future. This possibility was considered in a sensitivity analysis. The attribute levels for test effectiveness and reliability were based on the results of the retrospective sample analysis, presented in Chapter Four.

Valuing the attribute levels for each comparator

The final step in the monetisation process is to value the different test specifications described in **Table 5.6**. Several choice models were presented in Section 5.4. Of these, the mixed logit model for all respondents (**Table 5.3**) provided the best fit, hence the WTP values from this model were taken forward for the CBA.

5.5.2 Cost data used in the cost-benefit analysis

The key consideration when estimating costs in the CBA is that the timeframe over which the costs are accrued should match the timeframe that patients were asked to consider when making their choices in the DCE [86]. This timeframe was not explicitly stated in the DCE. However, the aim of the DCE was to establish the value that patients place on possessing the information provided by genetic and genomic testing. It therefore seems reasonable to assume that the appropriate timeframe is equivalent to the time between undergoing the genetic or genomic test and getting the test results (specified in **Table 5.6** as 14 days). Within this timeframe, the costs potentially incurred by patients include the cost of genetic or genomic testing and the costs of any resource use related to undergoing testing (from both a healthcare provider and societal perspective). As the entire testing process takes place during existing appointments and does not result in any additional tests or procedures, costs in the latter category were determined to be zero. In the former category, genetic testing cost £281 while genomic testing cost £319 (**Appendix 16**).

If the WTP values from the DCE did represent patient WTP for both testing and treatment, it would then be possible to compare these values with the mean cost of each comparator in the CEA and CUA. As this is not possible, an alternative approach – benefit transfer

analysis – was considered as an extension to the CBA presented in this section. This approach estimates benefits in one context by adapting estimates of benefits from another context and is described in Chapter Six.

5.5.3 Data analysis

The final step in the CBA process is to combine the data on costs and benefits to calculate the net benefit of both current genetic testing practice and future genomic testing practice. If net benefit is positive for a particular form of testing, this suggests that it is worthwhile to adopt this intervention. If net benefit is positive for both forms of testing, the test with the highest net benefit would normally be recommended for adoption.

The uncertainty surrounding the net benefit estimates was evaluated in several ways. First, a comprehensive univariate sensitivity analysis was conducted in which the WTP estimates for all attributes were varied between their 95% confidence intervals. Other variations included changes in the cost of genetic and genomic testing, improvements in the efficacy and speed of genomic testing, changes in type of staff and differing appointment lengths. Second, the coefficient values from the mixed logit model were used to conduct a probabilistic sensitivity analysis (PSA) with 1,000 simulations, with normal distributions attached to attributes, as per a previous study [25]. This permitted the calculation of 95% confidence intervals around the net benefit estimates for both forms of testing, as well as for the net benefit difference between the two tests. The proportion of instances in which genetic (or genomic) testing yielded the highest net benefit in the PSA was also calculated. Finally, the impact of risk graphics on net benefit was evaluated by applying WTP values from the mixed logit models for respondents who did (or did not) view risk graphics.

5.6 Cost-benefit analysis results

5.6.1 Main results

The CBA results are presented in **Table 5.7**. WTP for genetic testing is -£2,960 while WTP for genomic testing is -£1,798, a difference of £1,162. The negative WTP estimates are primarily driven by the WTP for improvements in test reliability. As genetic testing costs £281 per patient, the overall net benefit associated with genetic testing is -£3,241. Genomic testing is more expensive (£319), but as WTP is higher, net benefit is -£2,117. The difference in net benefit between the two approaches is £1,124.

Table 5.7: Main cost-benefit analysis results

Attribute	WTP	Attribute levels		WTP	
		Genetic test	Genomic test	Genetic test	Genomic test
TIME	-£31.05	14 days	14 days	-£434.73	-£434.73
EFFECT	£53.79	6.77%	16.67%	£364.00	£896.53
REL	-£155.73	22.22%	18.18%	-£3,460.77	-£2,831.54
INFO0	-£60.66	-	-	£0.00	£0.00
INFO1	-£29.67	-	-	£0.00	£0.00
INFO2	£84.31	15 minute appointment	15 minute appointment	£84.31	£84.31
INFO3	£6.02	-	-	£0.00	£0.00
WHO0	-£543.81	-	-	£0.00	£0.00
WHO1	£265.58	-	-	£0.00	£0.00
WHO2	-£209.22	-	-	£0.00	£0.00
WHO3	£487.46	Consultant hospital doctor	Consultant hospital doctor	£487.46	£487.46
TOTAL WTP				-£2,959.73	-£1,797.97
TEST COST				£281.08	£319.23
NET BENEFIT				-£3,240.81	-£2,117.20
DIFFERENCE				-	+£1,123.60

WTP = willingness-to-pay. TIME = time to receive test result; EFFECT = ability of the test to predict who will not respond to the usual chemotherapy treatment; REL = test reliability; INFO = length of time clinicians spend describing the test to you (0=5min, 1=10min, 2=15min, 3=20min); WHO = type of clinician who explains the test result to you (0=general practitioner, 1= specialist nurse, 2=junior hospital doctor, 3=consultant hospital doctor).

5.6.2 Sensitivity analysis and probabilistic sensitivity analysis

Appendix 35 presents the results of the sensitivity analysis. There were no parameter variations which changed the finding that genomic testing has the highest net benefit. However, several variations did produce notable changes in net benefit, including an increase in WTP for reliability improvements from -£156 to -£91. This increased WTP for both forms of testing, but as genetic testing is less reliable than genomic testing in this context the net benefit difference between the two tests decreased. A second notable variation was the use of WTP values from the risk graphics model rather than the pooled model for all respondents. Again, WTP for both tests increased, but the increase for genetic testing was relatively greater, reducing the net benefit difference between the two tests. In addition, a scenario was considered in which the effectiveness and reliability of genomic testing were improved to simulate likely future improvements in test performance. This resulted in an increase in the net benefit of genomic testing from -£2,117 to +£1,007, and an overall increase in the net benefit difference between the two tests from £1,124 to £4,248. Finally, as the negative WTP estimates for genomic testing were primarily driven by test reliability, the level of reliability required to generate a positive net benefit for genomic testing was estimated. With genomic test effectiveness held constant, test reliability would need to improve to 4.59% for this to hold (i.e. out of every 22 patients identified as being a likely non-responder, only one of these is incorrectly identified).

The results of the PSA allowed 95% confidence intervals to be calculated for the net benefit estimates. The confidence intervals for genetic testing were -£5,141 to -£2,190 (base case: -£3,241), those for genomic testing were -£3,409 to -£1,332 (base case -£2,117) and those for the net benefit difference were £779 to £1,785 (base case £1,124). In all simulations,

genomic testing had the highest net benefit, as indicated in **Figure 5.2**. However, the highest net benefit estimate was never positive (**Figure 5.3**).

Figure 5.2: Net benefit differences between genetic and genomic testing

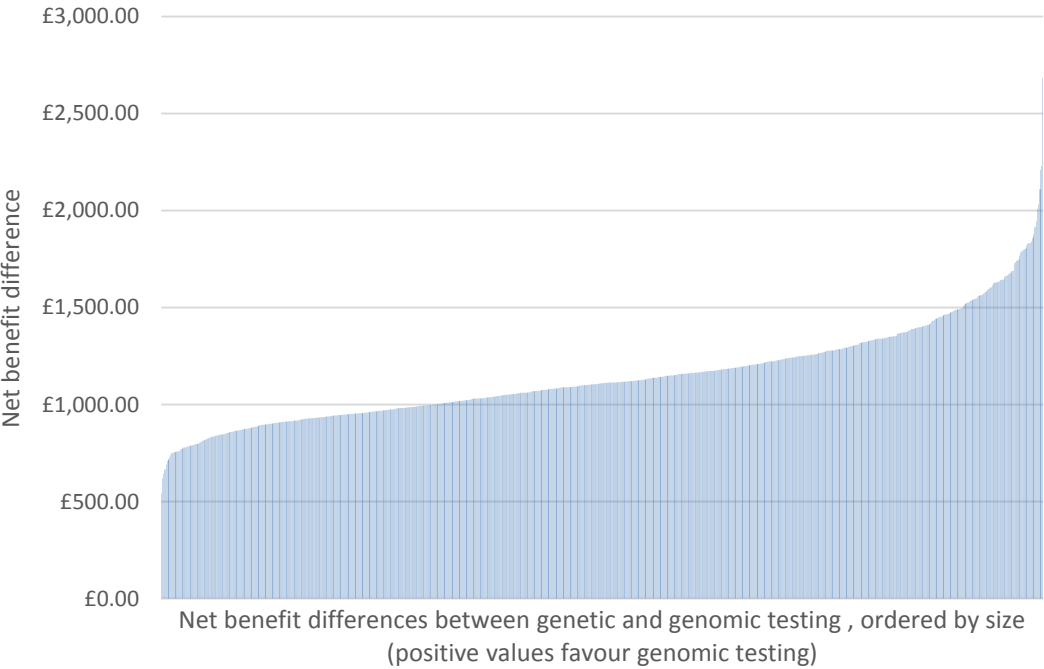
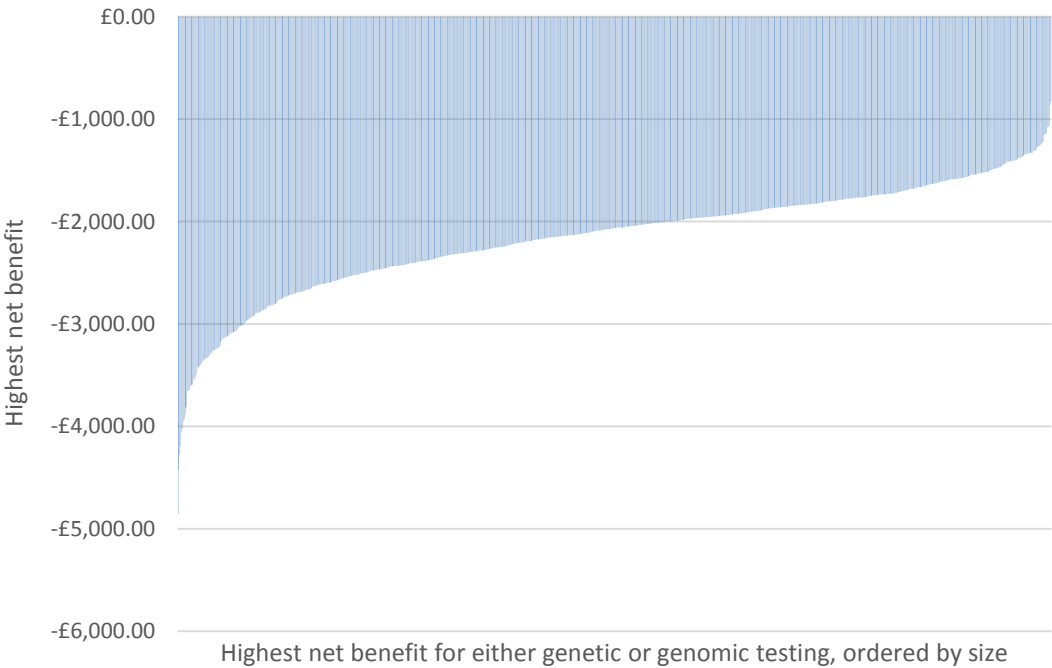


Figure 5.3: Highest net benefit for either genetic or genomic testing



5.6.3 Discussion

The main result from the CBA is that genomic testing has a higher net benefit than genetic testing, although net benefit is negative for both forms of testing. The univariate and probabilistic sensitivity analyses indicated little uncertainty around this result, although future scenarios were identified in which genomic testing may yield a positive net benefit.

To place the estimates of WTP for genetic and genomic testing into context it is useful to consider estimates from other DCEs conducted in similar contexts. The literature search documented in **Appendix 27** identified two studies that yielded such estimates. Chan *et al.* evaluated patient preferences for pharmacogenetic testing to guide warfarin treatment, with WTP varying between S\$340-S\$685 (£160-£321 at 2015 exchange rates) for the best-case test [257]. Regier *et al.* evaluated patient preferences for the use of array genomic hybridisation testing to identify children with idiopathic developmental disability, finding that WTP for the array test was CDN\$1118 (£552 at 2015 exchange rates) [258]. WTP for genetic and genomic testing in a CLL context is much lower than these two estimates. One possible explanation for this result is that test reliability is very important for CLL patients: they value a test which doesn't misidentify patients as non-responders because the consequences of doing so are perceived to be undesirable (patients don't receive FCR – known to be effective – to which they would likely have responded). For tests that primarily provide information, such as the use of array genomic hybridisation considered by Regier *et al.*, it is possible that test reliability is less important as there are fewer negative health effects associated with an incorrect result, making a positive WTP estimate more likely. A second possible explanation relates to the levels used for the reliability attribute in the DCE. This attribute describes how many times FCR non-responders are identified

incorrectly in every 100 tests, and the levels used for this attribute in the DCE were 2, 4, 6 and 8 times out of 100. After the DCE was conducted it became apparent that the reliability of genetic and genomic testing differed from the levels used in the choices in the DCE, the actual values being 22 and 18 out of 100, respectively. The high negative WTP value attached to the reliability of genetic and genomic testing may therefore be driven by the extrapolation of WTP estimates beyond the levels considered in the DCE.

The differences in WTP when risk graphics are used (compared to no graphics) raise some interesting questions regarding the true WTP of patients for genetic and genomic testing in this context. In Section 5.2.3 it was noted that if respondents misinterpret risk information in DCEs, this could invalidate WTP estimates, leading to suboptimal decision-making. It was then suggested that risk graphics may enhance the sensitivity of respondents to the magnitude of a risk, better capturing the true value of an intervention and potentially improving decision-making. Although the results presented in this chapter quite convincingly indicate that the preferences of respondents who are presented with risk information using graphics differ from those of respondents who are presented with the same information without graphics, it is difficult to argue that this difference reflects an improvement in risk comprehension without additional information on how respondents made their choices. In future studies, this additional information could be derived from qualitative work (e.g. think aloud studies, or investigations into whether risk graphics impact on survey fatigue) or from behavioural studies undertaken alongside a DCE (e.g. using eye-tracking software to gather data on how long respondents view choice questions both with and without risk graphics).

Furthermore, although there is a difference in WTP for genomic testing by use of risk graphics, it cannot necessarily be stated that this indicates improved risk comprehension. To do so, 'gold standard' information is required on the true WTP of these patients. In the longer term, such information may be available from revealed preference studies conducted to evaluate WTP for direct-to-consumer genetic tests (e.g. the 23andMe test: www.23andme.com). For now, a first step towards identifying the true WTP of patients might be to conduct stated preference studies using multiple approaches simultaneously, in order to 'triangulate' WTP estimates and thus generate a gold standard estimate.

Drummond *et al.* suggest that in instances where actual markets do not exist to validate WTP findings, it can be useful to consider the construct validity of results i.e. are the data consistent with theoretical constructs that should be present if the WTP responses are measuring the intended value? A key 'construct' that could be tested is that *"the more of a positively valued good that is supplied by a hypothetical programme, the greater should be a persons' WTP"* [12]. This 'scope test' is satisfied in both the overall DCE model and the two models in which use of risk graphics varies, although there are some differences in the magnitude of WTP. Consequently this test does not assist with identifying which of the WTP estimates represents the true WTP of patients in this context.

5.7 Conclusions

This chapter has described the design and results of a DCE which elicited information on the preferences of CLL patients for the use of genetic or genomic testing to guide chemotherapy treatment decisions. The results of this DCE suggest that respondents prefer combinations of attributes and attribute levels that more closely reflect genomic testing

than genetic testing. This finding was reinforced by the overall estimates of WTP for genetic and genomic testing, with the highest WTP value attached to genomic testing. However, the preferences of respondents who are presented with genetic risk information using graphics differ from those of respondents who are presented with the same information without risk graphics. These DCE results were then used within a CBA which considered the costs and consequences of genetic and genomic testing from a welfarist perspective. The results of this CBA indicated that the net benefit associated with genomic testing is higher than that for genetic testing. There was also some evidence to suggest that different risk presentation approaches could impact on this result.

These results have implications for decision-making in both CLL and genomics more generally. CLL patients clearly value the additional information that could be provided by genomic tests, however, several process-related attributes are also important to them. This suggests that for genomic testing to be implemented successfully, commissioners will need to ensure that test results are delivered by approved CLL clinicians in an appropriate environment. Furthermore, there is evidence to suggest that implementation may need to be tailored to the preferences of specific patient subgroups, with gender, income and previous treatment experience all influencing preferences.

Chapters Four and Five have described two analyses conducted to address the primary research question of this thesis: what is the most appropriate economic evaluation approach to use when assessing genomic interventions? The results of these analyses differ in several ways, and Chapter Six presents a comparison of these two approaches to economic evaluation in the context of genetic and genomic testing in CLL.

Chapter Six | Comparing the welfarist and extra-welfarist approaches to economic evaluation in chronic lymphocytic leukaemia

6.1 Introduction

Chapters Four and Five described a welfarist and an extra-welfarist economic evaluation, conducted to determine the most appropriate economic evaluation approach to use when assessing genomic interventions. The results of these analyses differed and it is important to compare them in a transparent, objective and thorough manner before identifying the most appropriate approach in this context.

Such a comparison is complicated. In Chapter Two it was noted that the two approaches are believed to ask fundamentally different resource allocation questions and are therefore unlikely to be equivalent [14,58]. This means that analysts should not simply convert extra-welfarist economic evaluations into welfarist economic evaluations and then directly compare the results; a more sophisticated approach is required. In the context of this thesis, this necessitated adjustments to the analytical approaches used in both forms of economic evaluation, and some deviations from the standard 'reference case' specified by health technology assessment (HTA) agencies such as the National Institute for Health and Care Excellence (NICE), including extending the analysis from the recommended national health service (NHS) perspective to a societal perspective [13]. A full description of these methodological deviations and their consequences is provided in Section 6.2.

The additional quantitative evidence generated in these extended analyses facilitated a more objective comparison of the results of the welfarist and extra-welfarist economic evaluations, and Section 6.3 describes this comparison. Both the adoption decisions suggested by each form of economic evaluation and the costs associated with reducing any uncertainty present in these adoption decisions are considered. Methodological issues related to the outcome measures used in each approach are also highlighted.

Quantitative evidence is useful in this context, but only to a point: ultimately this evidence is interpreted by healthcare decision-makers prior to reaching an adoption decision. Knowledge of the views of these stakeholders would therefore also be useful when comparing the welfarist and extra-welfarist approaches. Section 6.4 provides an insight into these views by describing a stakeholder survey which aimed to assess the degree to which the different economic evaluation approaches captured the true costs and benefits of genomic testing in chronic lymphocytic leukaemia (CLL). The chapter then concludes with a summary of the evidence that has been presented.

6.2 Facilitating a quantitative comparison between the two approaches

Four analytical extensions were required to facilitate a quantitative comparison between the two economic evaluation approaches. The following sub-sections explain how these extensions enabled such a comparison, describe the revisions that were undertaken and report the results for each analytical variation.

6.2.1 Extending the extra-welfarist economic evaluation to a societal perspective

The costing perspective selected for the extra-welfarist economic evaluation in Chapter Four was that of the NHS, hence only the costs incurred by the NHS when treating CLL patients were included, as per NICE guidance [13]. However, it was noted that economic evaluations of genomic tests may require a societal costing perspective to be adopted as such interventions may have impacts beyond the perspective of the healthcare sector. This broader perspective is commonly adopted in welfarist economic evaluations as the aim of these studies is considered to be to maximise societal welfare with reference to a societal budget constraint [14]. As such, the welfarist economic evaluation in Chapter Five nominally adopted a societal perspective (although the timeframe over which costs were accrued in the cost-benefit analysis (CBA) was relatively short, hence this decision did not impact on the CBA results). Consequently, to facilitate a more objective comparison between the two sets of results, both the cost-effectiveness analysis (CEA) and the cost-utility analysis (CUA) were extended to consider a societal costing perspective.

Methods

Several references are available to guide analysts wishing to extend economic evaluations to a societal perspective [12,31,86,246,294]. There is a degree of consensus concerning which costs should be included in these analyses, with three main cost categories identified:

1. Productivity costs. These costs are incurred when the productivity of CLL patients is affected by illness, treatment, disability or premature death. Productivity costs fall into several sub-categories [92]:

- a. Short-term absenteeism due to CLL-related morbidity (e.g. time taken off work or leisure to attend routine healthcare appointments);
 - b. Short-term presenteeism due to CLL-related morbidity (presenteeism captures the impact of reduced productivity during work or leisure time);
 - c. Long-term absenteeism due to CLL-related morbidity;
 - d. Productivity losses due to CLL-related mortality;
 - e. Productivity gains due to improved treatment.
2. Informal care costs. CLL patients frequently need informal care in addition to the care provided by healthcare professionals. This category captures the costs of the time invested in patient care by informal unpaid caregivers, such as family or friends [92].
3. Out of pocket costs. These costs can be incurred by both CLL patients and informal caregivers and fall into several sub-categories, including:
- a. Costs associated with receiving outpatient and inpatient care (e.g. travel costs);
 - b. Costs associated with day-to-day living (e.g. home help);
 - c. Other healthcare costs (e.g. over the counter medicines);
 - d. Clothing and equipment costs (e.g. a raised toilet or hand rails).

Appendix 36 describes the methods used to estimate costs in each category. In brief, unit cost and resource use data were extracted from a variety of sources, including journal articles and reports from charities [295,296]. CLL-specific sources in a UK setting were favoured, but these sources are rare, hence data were also sourced from non-UK studies and studies in other diseases. Several assumptions were made in order to estimate the additional resources used from a societal perspective in different model health states. For example, it was assumed that days in hospital could only be experienced during

symptomatic model states, while physician contacts were assumed to require half a day away from employment. The costs associated with both long-term absenteeism and productivity losses due to CLL-related mortality were included, but as the starting age of patients in the model is 65 years (i.e. retirement age), these costs had no effect in the base case analysis. Starting ages of 25 (the likely youngest age of presentation [163]) and 45 years were evaluated in a sensitivity analysis to consider the impact of including these costs. Productivity gains due to improved treatment were not included; it was assumed that these gains were captured by the utility scores used in the CUA.

The analysis from a societal perspective was conducted in the same way as the base case analysis. Societal costs were accrued by patients in specific model states, in addition to the costs accrued from a NHS perspective. All of the parameters that were used in the societal cost calculations were varied in univariate sensitivity analysis, and probabilistic sensitivity analysis (PSA) was also used to further consider the impact of parameter uncertainty, and to calculate cost-effectiveness acceptability curves (CEACs).

Results

The results of this analysis are summarised in **Table 6.1** and presented in full in **Appendix 36**. The increase in mean costs per patient for each comparator is minimal compared to the base case results, hence the same comparators are dominated (or extendedly dominated), the incremental cost-effectiveness ratios (ICERs) are approximately the same, and the results of the PSA suggest the same level of uncertainty around these results. The concordance between the CEA and CUA results that was observed in the base case analysis is therefore replicated when a societal perspective is considered.

Table 6.1: Results for the cost-effectiveness and cost-utility analyses from a societal perspective

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies) ^a	ICER (excluding extendedly dominated strategies) ^b
CEA	C	6.37	£73,832	-	-
	A	6.61	£76,035	£9,302	£9,302
	Int 2	6.65	£95,031	£545,428	EXT.DOM
	B	7.63	£110,820	£16,007	£34,062
	Int 1	7.45	£122,116	DOM	DOM
CUA	C	5.60	£73,832	-	-
	A	5.82	£76,035	£10,081	£10,081
	Int 2	5.93	£95,031	£166,523	EXT.DOM
	B	6.44	£110,820	£30,908	EXT.DOM
	Int 1	6.67	£122,116	£50,164	£54,207

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life-years; QALYs = quality-adjusted life years. ^a Dominated strategies are less effective than cheaper alternatives. ^b Extendedly dominated strategies have an ICER that is higher than the next most effective strategy.

The only variation which changes this result is the use of different starting ages. **Table 6.2** presents the results from a societal perspective for patients aged 25. Mean costs increase for all strategies because CLL patients now die at a younger age, increasing the costs associated with premature mortality. The main consequence of this cost increase in the CEA is that the most cost-effective strategy changes from Comparator A (current practice in the Oxford University Hospitals NHS Trust) to Comparator B (current practice with ibrutinib given as refractory treatment for all patients) with an ICER of £2,943 per life-year gained. The most cost-effective strategy in the CUA also changes, but to Intervention One (genomic testing followed by first-line ibrutinib treatment in likely FCR non-responders) with an ICER of £19,933 per quality-adjusted life year (QALY) gained. Hence, at younger ages, strategies which include ibrutinib treatment are likely to be the most cost-effective option from a societal perspective as these minimise premature mortality.

Table 6.2: Results for the cost-effectiveness and cost-utility analyses from a societal perspective for patients aged 25

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	Int 2	7.73	£687,062	-	-
	B	8.88	£690,438	£2,943	£2,943
	A	7.35	£697,523	DOM	DOM
	Int 1	8.73	£698,459	DOM	DOM
	C	7.00	£707,752	DOM	DOM
CUA	Int 2	6.91	£687,062	-	-
	B	7.42	£690,438	£6,701	£6,701
	A	6.47	£697,523	DOM	DOM
	Int 1	7.82	£698,459	£19,933	£19,933
	C	6.15	£707,752	DOM	DOM

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life-years; QALYs = quality-adjusted life years.

In addition, discordance between the results of the two forms of extra-welfarist economic evaluation is more likely at younger ages when a societal perspective is considered. This is because the CEA favours interventions which reduce mortality and the associated high societal costs (in this case, Comparator B: genetic testing and refractory ibrutinib treatment), whereas the CUA favours interventions which also improve quality of life (in this case, Intervention One: genomic testing stratifies patients more accurately to ibrutinib treatment, so fewer patients receive inappropriate treatment, and thus fewer patients experience adverse events, which impacts on absenteeism and presenteeism).

6.2.2 Extending the extra-welfarist economic evaluation to consider different time horizons

The second adjustment that was made to the extra-welfarist economic evaluation was to consider different time horizons. In the base case analysis a time horizon of 30 years was

selected. However, the time horizon considered in the CBA was much shorter at 14 days, informed by the likely time horizon that patients considered when making choices in the discrete choice experiment (DCE). As the cycle length in the model underlying the CEA and CUA was set at 28 days, it was not possible to calculate these results over a time horizon of 14 days, nor was it desirable to do so, as few costs or outcomes would be accrued in this short period. However, the results for other time periods were calculated to assess the impact of varying this assumption. Three alternative time horizons were considered – 20 years, ten years and one year – to test an extreme assumption (one year) and two horizons of average length. The analytical approach for the CEA and CUA remained the same as in the base case analysis, and this extension was only undertaken for the analysis which considered a NHS costing perspective as the results of the analysis from a societal perspective were very similar.

The results of this analysis are presented in **Tables 6.3-6.5**. As the time horizon shortens, patients accrue fewer life-years, QALYs and costs. For time horizons of ten and 20 years the most cost-effective strategy remains Comparator A. For a time horizon of one year, the most cost-effective strategy is Intervention One, primarily because this is the cheapest strategy (£27,625 per patient). The results of the CEA and CUA remain concordant regardless of the time horizon. These results suggest that over short time horizons (approximating the time horizon considered in the CBA) a strategy which includes genomic testing and ibrutinib treatment is the most cost-effective option. This minimises mortality in this shorter timeframe, and the additional costs that are accrued due to following such an approach are realised in later years, after patients have undergone longer periods of ibrutinib treatment. Taken together, these two effects result in lower ICERs.

Table 6.3: Results for the cost-effectiveness and cost-utility analyses for a 20 year time horizon

Analysis	Comparator	Mean LYs/QALYs per patient ^a	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.33	£69,317	-	-
	A	6.54	£71,045	£8,027	£8,027
	Int 2	6.49	£88,387	DOM	DOM
	B	7.46	£103,177	£35,264	£35,264
	Int 1	7.26	£114,671	DOM	DOM
CUA	C	5.56	£69,317	-	-
	A	5.76	£71,045	£8,670	£8,670
	Int 2	5.79	£88,387	£595,249	Ext Dom
	B	6.31	£103,177	£28,102	£57,850
	Int 1	6.50	£114,671	£62,635	£62,635

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life-years; QALYs = quality-adjusted life years. ^a The cycle length in the model is 28 days, hence it is not possible to calculate results for a time horizon of exactly 20 years. These results are therefore calculated for a time horizon of 7308 days (20.01 years).

Table 6.4: Results for the cost-effectiveness and cost-utility analyses for a ten year time horizon

Analysis	Comparator	Mean LYs/QALYs per patient ^a	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	5.60	£63,081		-
	A	5.67	£64,102	£15,778	£15,778
	Int 2	5.28	£69,395	DOM	DOM
	B	6.05	£76,658	£32,833	£32,833
	Int 1	5.84	£87,996	DOM	DOM
CUA	C	4.92	£63,081		-
	A	4.98	£64,102	£16,099	£16,099
	Int 2	4.70	£69,395	DOM	DOM
	B	5.21	£76,658	£54,578	£54,578
	Int 1	5.22	£87,996	£1,915,911	£1,915,911

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life-years; QALYs = quality-adjusted life years. ^a The cycle length in the model is 28 days, hence it is not possible to calculate results for a time horizon of exactly ten years. These results are therefore calculated for a time horizon of 3668 days (10.04 years).

Table 6.5: Results for the cost-effectiveness and cost-utility analyses for a one year time horizon

Analysis	Comparator	Mean LYs/QALYs per patient ^a	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	Int 1	1.06	£27,625	-	-
	C	1.08	£28,551	£46,574	£46,574
	Int 2	1.04	£30,098	DOM	DOM
	B	1.07	£30,263	DOM	DOM
	A	1.07	£30,954	DOM	DOM
CUA	Int 1	0.91	£27,625	-	-
	C	0.90	£28,551	DOM	DOM
	Int 2	0.88	£30,098	DOM	DOM
	B	0.90	£30,263	DOM	DOM
	A	0.90	£30,954	DOM	DOM

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life-years; QALYs = quality-adjusted life years. ^a The cycle length in the model is 28 days, hence it is not possible to calculate results for a time horizon of exactly one year. These results are therefore calculated for a time horizon of 392 days (1.07 years). Mean life-years per patient for Comparator C surpass the length of the time horizon due to the implementation of a half-cycle correction.

6.2.3 Undertaking value of information analysis

The results of the extra-welfarist economic evaluation from both costing perspectives indicated some uncertainty surrounding the ICERs that were calculated for non-dominated strategies. This uncertainty was captured by the CEACs: although Comparator A was the most cost-effective strategy at a threshold of £30,000 per life-year/QALY in both the CEA and CUA, the CEACs indicated that there was only a 65-66% probability of this conclusion being correct. This information provides important context for the results of the CEA and CUA and is therefore helpful when comparing the results of the different forms of economic evaluation. However, knowledge of the existence of uncertainty must be combined with estimates of the costs associated with reducing this uncertainty for this information to be of most use when comparing different economic evaluation approaches. If, for example, there is sufficient uncertainty to warrant delaying an adoption decision

until better evidence is available, knowledge that the cost of collecting such evidence is likely to be greater than the benefits may change an adoption decision from 'delay until additional research is undertaken' to 'do not adopt'. Given this, a value of information (VOI) analysis was conducted to provide context for the uncertainty in the CEA and CUA results.

Value of Information analysis background and methods

VOI analysis can inform decision-makers about the costs of uncertainty and the value of collecting further evidence to reduce this uncertainty. This is achieved by calculating the expected value of perfect information (EVPI), which is an estimate of the value of future research which would eliminate the uncertainty surrounding all model parameters. EVPI is simply the difference between the expected value of making an adoption decision based on current information, and the expected value of making the same decision with perfect information. VOI analysis has been shown to be useful when addressing a variety of policy questions [297], and NICE currently recommends its use when evaluating public health interventions [67].

Calculation of EVPI is based on the PSA results already calculated for the extra-welfarist economic evaluation and requires two further pieces of information: the effective population, and the lifetime of the intervention. The effective population is the annual population of UK CLL patients who would undergo genetic or genomic testing, and the lifetime of the intervention is the period over which the EVPI analysis is conducted (not necessarily the same as the time horizons evaluated in Section 6.2.2). Data on the effective population for genetic or genomic testing in a UK context is not available, so this was informed by a study which evaluated the burden of CLL within the EU [298]. This study

estimated that 5,581 CLL patients would require treatment in a typical year, and the implied incidence rate was used to calculate an annual effective population of 719 for the UK population in 2013 [299]. The lifetime of the intervention is uncertain and thus more difficult to estimate. EVPI was therefore calculated for three time horizons (ten, 15 and 20 years) for both the CEA and CUA, and from both a NHS and a societal costing perspective.

Results

Table 6.6 presents the EVPI results for different threshold ICERs. **Appendix 37** presents these results graphically. At a threshold of £30,000 per life-year/QALY, EVPI varies between £15m (CEA, NHS perspective, ten year time horizon) and £30m (CUA, societal perspective, 20 year time horizon). EVPI always increases as the time horizon increases, and the results from a NHS and societal perspective are very similar (although the value of EVPI is always slightly higher from a societal perspective, as there is slightly more uncertainty surrounding these results). This range of EVPI values (£15m-£30m) represents the upper limit of the cost of research that should be considered to reduce uncertainty for this decision problem.

Table 6.6: Population expected value of perfect information

Threshold ICER (£)	Population EVPI (£m, 2013 costs)											
	NHS perspective						Societal perspective					
	10 years		15 years		20 years		10 years		15 years		20 years	
	CEA	CUA	CEA	CUA	CEA	CUA	CEA	CUA	CEA	CUA	CEA	CUA
£5,000	£5m	£5m	£7m	£7m	£9m	£8m	£5m	£4m	£7m	£6m	£8m	£8m
£10,000	£8m	£9m	£12m	£12m	£14m	£15m	£10m	£9m	£14m	£13m	£17m	£16m
£15,000	£9m	£9m	£12m	£12m	£15m	£15m	£10m	£10m	£14m	£14m	£17m	£17m
£20,000	£10m	£10m	£13m	£14m	£17m	£17m	£11m	£11m	£15m	£16m	£19m	£19m
£25,000	£12m	£12m	£17m	£16m	£20m	£20m	£14m	£14m	£19m	£19m	£23m	£23m
£30,000	£15m	£15m	£21m	£21m	£25m	£25m	£17m	£17m	£24m	£24m	£29m	£30m
£35,000	£20m	£19m	£28m	£26m	£34m	£32m	£22m	£22m	£30m	£31m	£37m	£38m
£40,000	£16m	£25m	£23m	£35m	£28m	£43m	£17m	£30m	£23m	£42m	£29m	£52m
£45,000	£15m	£35m	£20m	£48m	£25m	£59m	£16m	£42m	£23m	£58m	£28m	£71m
£50,000	£15m	£47m	£20m	£65m	£25m	£81m	£17m	£57m	£23m	£79m	£29m	£97m

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; EVPI = expected value of perfect information; ICER = incremental cost-effectiveness ratio; NHS = National Health Service.

To place this figure into context, estimates of EVPI in other studies are required, along with information on whether further research was judged to be worthwhile. A UK study previously collected such data across several disease areas [297]. EVPI for a number of interventions fell into a range of £15m-£30m, including screening for age-related macular degeneration (£25m) and physiotherapy for children with asthma (£15m). In all cases further research was judged to be worthwhile as EVPI exceeded the cost of this research. A further study considering the case for genetic screening for haemochromatosis in Germany reported EVPI values in the range €0.5m-€2.2m, noting that these estimates may not exceed the cost of the research required [153]. This range of values is, however, far lower than the values reported in this thesis. It is therefore likely that the benefits (in monetary terms) of research which reduces the uncertainty in the extra-welfarist economic evaluation results exceed the cost of the research. The exact nature of the research that may be required is discussed in Chapter Seven.

6.2.4 Adapting the welfarist economic evaluation

The welfarist economic evaluation was based on a DCE which aimed to establish the value that patients place on possessing the information provided by genetic and genomic testing. Consequently the CBA that was conducted compared the costs of testing – accrued over a relatively short time period – with patient willingness-to-pay (WTP) for the information provided by these tests. The narrow specification of this CBA is in contrast to the extra-welfarist economic evaluation which considered five comparators, reflecting different configurations of genetic and genomic testing followed by treatment. Therefore, in order to permit a more objective comparison of the results of two approaches, it was necessary to extend the CBA to consider WTP more broadly.

One way in which this could be achieved is by collecting data on the WTP of CLL patients for treatment following testing, combining this with the estimates of WTP for testing, then comparing this combined WTP value with the mean cost per patient of each of the strategies considered in the extra-welfarist economic evaluation. However, data on WTP for treatment following testing was not available from the study underlying this thesis. A second option is to conduct a benefit transfer analysis, in which benefits are estimated in one context by adapting estimates of benefits from another context [86]. In a non-health context, this could mean using an estimate of WTP for an ecosystem in one region to value a similar ecosystem in another region. For this study, benefit transfer would require the identification of estimates of WTP for treatment following genomic testing in a non-CLL context, then the adaptation of these estimates to reflect the use of genomic testing in CLL. However, such estimates are also not available in the literature, ruling out a conventional benefit transfer analysis.

Instead, the welfarist economic evaluation was extended by adapting the benefit transfer approach to utilise the available data in this context. In practical terms, the benefit transfer framework was used to estimate the value of WTP for treatment following testing that would be required for overall WTP for genomic testing (including WTP for testing alone) to surpass the cost of testing and treatment. The value elicited in this 'what if' analysis was then compared with estimates of WTP from the literature to determine the likelihood that this value would be obtained, were such a WTP study undertaken in a CLL context.

This extension focused on the two strategies in which genomic testing is used to guide treatment decisions – Interventions One and Two – in order to identify a WTP value which

might support adoption of these tests from a welfarist perspective. To calculate this value, information was required on WTP for genomic testing (provided by the DCE), the mean number of genomic tests per patient (extracted from the Markov model), the mean societal cost per patient (estimated in Section 6.2.1), and the mean life-year or QALY gain per patient (extracted from the Markov model). **Table 6.7** presents the results of this analysis.

Table 6.7: ‘What if’ benefit transfer analysis results

Calculation step ^a	Intervention One	Intervention Two
A. Mean life-years per patient	7.45	6.65
B. Mean QALYs per patient	6.67	5.93
C. Mean societal costs per patient	£122,116	£95,031
D. Willingness-to-pay for genomic testing	-£1,798	-£1,798
E. Mean number of genomic tests per patient	5.36	3.79
F. Total willingness-to-pay for genomic testing (D x E)	-£7,635	-£5,465
G. Willingness-to-pay gap (C – F)	£129,751	£100,496
H. Required willingness-to-pay for one life-year to yield a positive net benefit for genomic testing (G/A)	£17,481	£15,122
I. Required willingness-to-pay for one QALY to yield a positive net benefit for genomic testing (G/B)	£19,460	£16,943

QALYs = quality-adjusted life years. ^a The calculations were carried out as follows. First, WTP for genomic testing was multiplied by the mean number of tests per patient, and the resulting value was subtracted from mean societal costs per patient to identify the value of WTP for treatment following testing that would be required for overall WTP to surpass the cost of testing and treatment. Information on the mean life-years or QALYs gained per patient for each strategy was then used to express this WTP value as a WTP for one life-year or QALY.

These results indicate that WTP for one life-year would need to be between £15,122-£17,481 for a benefit transfer approach to yield a positive net benefit for genomic testing. The equivalent range expressed in QALYs was £16,943-£19,460. To determine the likelihood of such a value being obtained, a search was undertaken for other studies estimating patient WTP for survival gains (or QALY gains) for any cancer treatment. Two studies were identified. The first estimated the WTP of patients from Singapore with advanced cancer to extend life by one year to be S\$18,570 per life-year gained (£8,728 in

2015 prices) [300]. The second study estimated the WTP of patients from the USA with metastatic prostate cancer for an additional life-year to be US\$1,170 (£749 in 2013 prices) [301]. These estimates are much lower than the values required to yield a positive net benefit for genomic testing. It was therefore considered unlikely that the sum of WTP for testing and treatment following testing would be greater than the societal costs associated with the genomic testing strategies.

6.3 Comparing the results of the two approaches: quantitative evidence

Section 6.2 described several analytical extensions to facilitate a more objective comparison of the results of the welfarist and extra-welfarist economic evaluations. This next section compares the adoption decisions suggested by each analytical approach. Methodological issues related to the outcome measures used in each approach are also highlighted.

6.3.1 Comparison of adoption decisions suggested by each approach

Table 6.8 presents a summary of the adoption decisions with respect to genomic testing that would likely be made by decision-makers based on the results of the different forms of economic evaluation. It is assumed that decision-makers evaluate the results of the extra-welfarist economic evaluations against a threshold ICER of £30,000 per life-year/QALY gained, adopting the most cost-effective intervention below this threshold. For the welfarist economic evaluation, it is assumed that decision-makers would adopt the intervention with the highest positive net benefit.

Table 6.8: Summary of the adoption decisions suggested by each analytical approach

Type of economic evaluation	Analytical approach	Likely adoption decision with respect to genomic testing	Justification
CEA	Base case analysis (NHS perspective)	Do not adopt genomic testing	Comparator A is the most cost-effective option. CEACs suggest some uncertainty around this decision, but the strategy with the second highest probability of being cost-effective is also a genetic testing strategy.
	Societal perspective	Do not adopt genomic testing	Comparator A is the most cost-effective option. CEACs suggest some uncertainty around this decision, but the strategy with the second highest probability of being cost-effective is also a genetic testing strategy. Comparator B is the most cost-effective strategy for younger patients.
	Short time horizon	Adopt genomic testing	Intervention One is the most cost-effective strategy.
	VOI analysis	Delay decision until further research conducted	Quite likely that EVPI exceeds the cost of the research required to reduce the uncertainty in the CEA results.
CUA	Base case analysis (NHS perspective)	Do not adopt genomic testing	Comparator A is the most cost-effective option. CEACs suggest some uncertainty around this decision, but the strategy with the second highest probability of being cost-effective is also a genetic testing strategy.
	Societal perspective	Do not adopt genomic testing for all patients. Consider genomic testing for younger patients	Comparator A is the most cost-effective option. CEACs suggest some uncertainty around this decision, but the strategy with the second highest probability of being cost-effective is also a genetic testing strategy. Intervention One (i.e. a genomic testing strategy) is the most cost-effective strategy for younger patients.
	Short time horizon	Unclear	Intervention One is the most cost-effective strategy, but with an ICER greater than £30,000/QALY.

Type of economic evaluation	Analytical approach	Likely adoption decision with respect to genomic testing	Justification
	VOI analysis	Delay decision until further research conducted	Quite likely that EVPI exceeds the cost of the research required to reduce the uncertainty in the CUA results.
CBA	Base case analysis	Unclear	The net benefit associated with genomic testing is higher than that associated with genetic testing, and this is a significant difference. However, the net benefit for both forms of testing is negative, so the likely adoption decision is unclear.
	'What if' benefit transfer analysis	Do not adopt genomic testing	It is very unlikely that the sum of WTP for testing and WTP for treatment following testing would be greater than the societal costs associated with either genomic testing strategy.

CEA = cost-effectiveness analysis; CEAC = cost-effectiveness acceptability curve; CUA = cost-utility analysis; CBA = cost-benefit analysis; EVPI = expected value of perfect information; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life-years; NHS = National Health Service; QALYs = quality-adjusted life years; VOI = value of information; WTP = willingness-to-pay.

Ten adoption decisions are presented. Five of these decisions are 'do not adopt genomic testing' (in one case there is a possible exception for younger patients), two decisions are 'delay a decision until further research is conducted', two further decisions are unclear, and in the final case the decision is 'adopt genomic testing'. Overall, only one of the ten decisions definitively supports the adoption of genomic testing.

In terms of the CEA, there is generally concordance between the adoption decisions suggested by each analytical approach. The base case analysis suggests that Comparator A is the most cost-effective option, hence genomic testing should not be adopted. There is some uncertainty around this result, although the strategy with the second highest probability of being cost-effective is Comparator C (also a genetic testing strategy). When

a societal costing perspective is adopted a similar result is observed (although Comparator B is likely to be the most cost-effective strategy for younger patients). The VOI analysis indicates that further research is warranted to reduce the uncertainty surrounding this adoption decision, hence delaying a decision until such research is undertaken may be justified. The only analytical variation which suggests that genomic testing should be adopted is when short time horizons are considered in an attempt to match the timeframe considered in the CBA. However, it is unlikely that decision-makers would attach a great deal of validity to this result as few of the costs and outcomes that are important (from either a NHS or a societal perspective) are captured in such short time horizons.

A similar pattern of likely adoption decisions is observed for the CUA. The base case analysis suggests that Comparator A is the most cost-effective option, although again there is some uncertainty around this result. The analysis from a societal perspective also favours the status quo, although there is some evidence to suggest that Intervention One (a genomic testing strategy) is the most cost-effective strategy for younger patients. The VOI analysis again indicates that delaying an adoption decision until further research is undertaken may be justified. Finally, when very short time horizons are considered Intervention One is the most cost-effective strategy. However, this is with an ICER of £30,396/QALY gained compared to a do nothing approach, which is higher than the standard threshold ICER.

For both the CEA and CUA, a lower threshold ICER of £20,000 per life-year/QALY was also considered to represent a more conservative decision-maker. However, this lower threshold had minimal impact on the likely adoption decisions.

In terms of the CBA, the base case analysis results suggest that the net benefit associated with genomic testing is higher than that associated with genetic testing, and that this difference is statistically significant. However, the net benefit for both forms of testing is negative, which complicates the interpretation of these results. Two of the three DCE-led CBAs that were identified in Chapter Five also reported negative net benefits for some of the strategies considered in these economic evaluations [25,26]. Reflecting on these results, McCormack *et al.* note (in the context of hernia surgery) that *“The relevance of [the negative net benefit of some surgery options] to decision-making is unclear as this is not the same as a comparison with no treatment. Even if it was decided that a hernia should not be treated, some costs (e.g. the cost of emergency operations to treat strangulated hernias) and effects (e.g. risk of long-term pain) would be incurred”* [25]. Burr *et al.* also comment (in the context of surveillance for ocular hypertension) that *“...the net benefits for all of the pathways except for the ‘SOH hospital’ pathway were negative. This indicates that compared with a ‘no surveillance pathway’ these surveillance pathways would not be regarded as being worthwhile”* [26]. The results of the economic evaluation reported in this thesis have a number of similar characteristics to those reported by McCormack *et al.* In particular, it is difficult to calculate patient WTP for a scenario where there is no genetic or genomic testing to guide treatment decisions, due to the lack of data on patient WTP for treatment that was previously highlighted. It is possible that WTP for no genetic or genomic testing is such that the net benefit for this strategy is greater than that for genetic or genomic testing. Such a result would permit the negative net benefits for genetic and genomic testing to be interpreted as being inferior to the net benefit for no testing, and thus it could be concluded that no testing is the optimal approach. However, it is also possible that the net benefit associated with no testing is less than that associated with

genetic or genomic testing (or both). The adoption decision suggested by the CBA results reported in this thesis is therefore unclear. However, when the CBA results are incorporated into the 'what if' benefit transfer analysis, a likely welfarist adoption decision emerges. The results of this analysis suggest that genomic testing should not be adopted, as it is unlikely that the sum of WTP for testing and for treatment following testing would be greater than the societal costs associated with genomic testing.

A final consideration when comparing the adoption decisions suggested by each economic evaluation approach is whether these decisions are all concordant. This appears to be the case: only one of the ten decisions is in favour of genomic testing, and this relies on somewhat unlikely restrictions (with respect to the time horizon) being placed on the extra-welfarist economic evaluations.

6.3.2 Methodological issues related to the outcome measures used in each approach

Consideration of the adoption decisions suggested by the welfarist and extra-welfarist economic evaluations is a useful first step in comparing the results of the two approaches. However, it is also important to take a closer look at any methodological issues related to the outcome measures used in each approach as it is predominantly differences in the measurement of outcomes that separate the different forms of economic evaluation. This section focuses on two key issues for each form of economic evaluation: the validity of the data sources from which outcomes data were extracted, and the likelihood that these data captured those elements of outcomes that are important in this context.

Cost-effectiveness analysis

The outcome measure used in the CEA was life-years, and the number of life-years accrued by patients in each strategy was calculated within the Markov model. This outcome measure was therefore directly informed by the transition probabilities which determined the rate at which patients moved into death states in the model. Good practice guidance for conducting model-based economic evaluations requires that these probabilities are ideally sourced from randomised controlled trials (RCTs) which directly compare the technology under appraisal with relevant comparators. These guidelines also note that *“such evidence may not always be available and may not be sufficient to quantify the effect of treatment over the course of the disease. Therefore, data from non-randomised studies may be required to supplement RCT data”* [13]. The transition probabilities that informed this Markov model were estimated using a variety of randomised and non-randomised data sources, including RCTs [302], phase II trials [303] and phase 1b-2 studies [201]. These sources had a number of limitations including short time horizons, few direct comparisons between the treatments modelled in the CEA, and limited data on treatment effects by genetic mutation status. This latter limitation is a concern (and may also be common to other disease areas), however the data that are available have been collected in studies with relatively large sample sizes compared to the overall CLL population [219,230], and it is unlikely that data will be available in the future to facilitate the robust estimation of treatment effects by genetic mutation status. Full details of these data sources are provided in **Appendix 14**. Due to these limitations, data were not formally synthesised to estimate transition probabilities, which may contribute to the uncertainty in the CEA results. This is a weakness of the outcome measure used in the CEA.

A second concern is whether life-years capture elements of outcomes that are important in this context. Evidence suggests that CLL shortens life expectancy in all age groups, hence mortality outcomes will be important for CLL patients [304]. However, CLL is a chronic disease: patients may live with the disease and its symptoms for many years. Patients may also suffer from complications following chemotherapy treatment. Quality of life is therefore also important to CLL patients [305]. Consequently an outcome measure which disregards quality of life is unlikely to reflect the outcomes that are important to patients and decision-makers.

Cost-utility analysis

The outcome measure used in the CUA was QALYs, and the number of QALYs accrued by patients in the Markov model was calculated by applying utility weights to estimates of life-years gained. Good practice guidance for estimating utility weights requires that changes in quality of life are reported directly by patients, with the utility of these changes based on public preferences obtained via a choice-based method (preferably the EuroQol five dimensions (EQ-5D) questionnaire) [13]. The utility weights that were used in the CUA are presented in **Table 4.4**. As the evidence base for CLL-specific outcomes data of this type is limited, utilities were primarily extracted from a study by Beusterien *et al.* which used a standard gamble approach to elicit preferences from the general public for health states associated with CLL treatment [239]. This approach does not strictly adhere to NICE guidelines. The utility weights were also provided by line of treatment, so were not treatment-specific. Beusterien *et al.* note these concerns, and suggest that an alternative approach is valid in this instance because generic health status measures such as the EQ-5D “*may not capture key psychological impacts associated with CLL treatment, for example,*

the impact of knowing that one is responding to treatment” (the new five-level version of the EQ-5D may be more sensitive than the three-level version, hence this statement may now be somewhat less valid [306]). The authors also note that HTA agencies may be willing to accept data from other sources if adequate justification for doing so is provided. Although there are some concerns regarding the validity of the source of utility data, in the context of a limited evidence base it is possible that decision-makers would accept the use of these utility weights.

A more pertinent question is whether QALYs are an appropriate outcome measure in CLL. Although QALYs capture the quality of life data that are missing when life-years are used, there may be other outcomes that are important to CLL patients that are not captured. Beusterien *et al.* note the importance of ‘knowing’ in this context [239], and other non-health outcomes that may be important to patients include possessing information on prognosis and process-related benefits such as improvements in waiting time. Beusterien *et al.* argue that their approach captures the value of ‘hope’ in this context, but it is not clear that other non-health benefits arising from the use of genomic testing to guide CLL treatment would be captured by QALYs. Furthermore, it is important to note that although non-health outcomes may be important to patients, they may be less important to decision-makers who are making adoption decisions with the aim of maximising health effects in a resource-constrained health system. Thus there is a conflict in CLL between what is appropriate from a patient perspective and a decision-maker perspective.

Cost-benefit analysis

The benefits associated with genetic and genomic testing were monetised in the CBA using a DCE. Several factors should be considered when evaluating the validity of this approach. First, as recommended in the literature, the DCE was conducted in a population with experience in the area of interest (UK CLL patients) [273]. The characteristics of the patients who responded were typical of CLL patients more generally, although the survey response rate was only 20%. The preferences elicited from respondents were generally rational, in line with prior expectations and similar to those reported by other genetic DCEs, which is another positive result. However, the use of these preferences to inform the CBA can be questioned for two reasons. First, the preferences of respondents who were presented with risk graphics differed from those of respondents who were not presented with graphics. The use of pooled estimates to calculate WTP for genetic and genomic testing may therefore be queried (although as it is not clear which group of respondents best understood the risk information, it is difficult to specify the precise impact of pooled estimates on WTP). Second, after the DCE was conducted it became apparent that the reliability of genetic and genomic testing differed from the information given to DCE respondents. WTP for improvements in reliability (and thus WTP for genetic and genomic testing) may therefore differ from that reported in the DCE.

A final consideration is whether this DCE is an appropriate source of outcomes data in this context. The DCE asked patients to make choices between different genetic or genomic tests based on information about both the health and non-health attributes of testing. It therefore seems reasonable to conclude that the preferences elicited from respondents provide information about both health and non-health outcomes, hence it is likely that this

approach would be considered more appropriate than the use of life-years or QALYs from a patient perspective (although possibly not from the perspective of a decision-maker focused on maximising health effects). However, an issue related to the design of this DCE has already been noted: no information is provided on WTP for treatment following testing. The health effects arising from such treatment will be an important outcome for both patients and decision-makers, hence it is possible that neither group would consider the CBA to be an appropriate source of outcomes data in this context.

6.3.3 Summary of the quantitative comparison

In this section, the adoption decisions suggested by the welfarist and extra-welfarist approaches have been compared and methodological issues related to the measurement of outcomes have also been evaluated. A case could be made for either CEA, CUA or CBA being the most appropriate approach in this context. This is not necessarily a surprising result. Although the welfarist and extra-welfarist approaches have been compared in a transparent and thorough manner, they represent alternative paradigms, hence choice of approach will be somewhat dependent on the objectives and opinions of the stakeholders making this decision. A comparison of the different economic evaluation approaches in genomics is therefore incomplete unless the views of key stakeholders are also considered. Section 6.4 presents the results of a survey which aimed to address this evidence gap.

6.4 Comparing the results of the two approaches: stakeholder survey

This section provides an overview of a survey of stakeholders that was conducted to assist with the identification of the most appropriate economic evaluation approach to use when assessing genomic interventions. Full details of this survey are provided in **Appendix 38**.

6.4.1 Survey methods

The main aim of the survey was to ask respondents which economic evaluation results captured the true costs and benefits of genomic testing in CLL. Secondary aims were to establish a consensus regarding an adoption decision, and to ask respondents whether standard HTA methods should be adapted in a genomic context. Target respondents included health economists in commissioning or policy-making positions, working in genomics or with a CBA background. Individuals with relevant backgrounds in genomics, policymaking, science research and medicine were also targeted. Respondents were identified by accessing relevant directories, literature reviews, HTA agency websites and by asking existing respondents to recommend additional respondents.

The first step in the survey design process was the creation of a document which summarised the economic evaluation results. A copy of this document is provided in **Appendix 39**. The second step was to develop a series of questions based on this document which reflected the primary and secondary aims of this survey. These questions were piloted in a group of health economists before being finalised. The final list of questions is provided in **Appendix 40**. Data were also collected on respondent location, occupation and prior experience of conducting economic evaluations.

The survey was conducted from July to August 2015 using the LimeSurvey online survey application [194]. 91 potential respondents were contacted by email to request participation and asked to complete the survey within one month. Non-responders received one reminder email. The survey data were analysed in Microsoft Excel and Stata 12 [240,277]. Responses were evaluated for the whole group and for relevant sub-groups,

and responses to open-ended questions were reviewed to identify common themes. After the survey had been conducted an error was discovered in the CBA results that were presented to survey respondents. The impact of this error on the validity of the survey results is discussed in **Appendix 38**.

6.4.2 Survey results

Thirty-eight responses were received (response rate 42%). A complete list of responses is provided in **Appendix 41**. Most respondents (37%) were from the UK and were health economists (82%). Most (71%) had been lead health economists on one or more CEAs or CUAs. Only 21% of respondents had been lead health economists on one or more CBAs, with most (61%) having limited or no experience of welfarist approaches.

Respondents were asked to make four decisions. Their responses are summarised in **Table 6.9**, and detailed results are presented in **Appendix 38**. First, respondents were asked which approaches best captured the costs and benefits of genomic testing in CLL. Most selected CUA from a societal perspective (39%) or from a NHS perspective (34%). CBA was the least selected approach (13%). A quarter of respondents (26%) said that they did not have enough information to make a decision. Those who favoured CUA noted that this was the most transparent and comprehensive approach, and the best approach to prioritise investments from a fixed budget. One respondent selected this approach as CBA did not include treatment effects and CEA did not include quality effects. CBA was criticised for being difficult to interpret, and even respondents who thought this approach was appropriate noted that it may not be the best approach for informing decision-making. Several respondents selected multiple approaches on the basis that in combination, this

would allow a reasonable characterisation of intervention benefits. Those who were unable to make a decision required more information on the objectives of the decision-maker in question, and on the budget that the decision-maker was allocating.

Table 6.9: Summary of the decisions made by respondents in the survey

Question	Response	% ^a
Which approach(es) best capture the costs and benefits of genomic testing in CLL? ^b	CEA (NHS)	24%
	CEA (Societal)	18%
	CUA (NHS)	34%
	CUA (Societal)	39%
	CBA	13%
	Lack of info ^c	26%
	None ^d	13%
What adoption decision would you recommend? ^e	Adopt	16%
	Adopt with conditions	18%
	Delay decision	45%
	Do not adopt	21%
Would it be helpful for healthcare decision-makers to consider CBA results alongside CEA or CUA results when determining whether to adopt genomic testing for CLL patients?	Yes	26%
	No	29%
	Unsure	45%
Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic and genomic tests?	Yes	55%
	Only in specific circumstances	16%
	No	29%

CBA = cost-benefit analysis; CEA = cost-effectiveness analysis; CLL = chronic lymphocytic leukaemia; CUA = cost-utility analysis; HTA = health technology assessment; NHS = National Health Service perspective. ^a N = 38. ^b Respondents could select more than one approach, hence these percentages do not sum to 100%. ^c The exact wording for this answer was: "I have not been provided with enough information to answer this question". ^d The exact wording for this answer was: "I don't believe that any of these results capture the true costs and benefits of genomic testing in CLL". ^e The exact wording for the four possible responses to this question was: 1. Adopt genomic testing; 2. Adopt genomic testing, but with some conditions e.g. for specific patient sub-groups, only after later lines of treatment; 3. Delay a decision until additional research is undertaken, and 4. Do not adopt genomic testing.

Second, respondents were asked what adoption decision should be made regarding genomic testing in CLL. Most (45%) would delay a decision because insufficient information was available. Respondents to a stakeholder survey in the USA came to a similar conclusion regarding the use of Oncotype DX in breast cancer [125]. Others would delay until the price

of ibrutinib falls or until further information is available on genomic testing in sub-groups. Those who would not adopt (21%) noted that comparators which included genetic testing offered better value for money. However, some respondents suggested that conditional adoption could be feasible alongside prospective collection of better data.

Third, respondents were asked if it would be helpful for healthcare decision-makers to consider CBA results alongside CEA or CUA results when determining whether to adopt genomic testing in CLL. Respondents were generally uncertain about whether this would be helpful. Those who found the CBA results to be useful noted the importance of considering the national perspective and incorporating evidence on real decision-making. Capturing the 'value of knowing' was also considered important. Respondents who thought decision-makers should not consider CBA results in this context focused on the advantages of the status quo, noting that most of the time CUAs (if done well) are sufficient. Others highlighted perceived deficiencies in the welfarist approach, stating that methods for determining WTP are not sufficiently reliable. A final group of respondents noted that welfarist approaches introduce uncertainty over interpretation and which result is believed, and place too much responsibility on the decision-maker to select an appropriate evaluative framework.

A follow-up question asked respondents to take a broader view of the use of welfarist approaches in health care. Several decision contexts were identified in which these approaches would likely be useful, including evaluation of rare diseases and clinical guidelines. However, it was noted that if welfarist approaches are considered in genomics, they should be considered in all contexts.

Finally, respondents were asked if standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic and genomic tests. Most (71%) replied 'yes', or 'yes in specific circumstances'. Key challenges that were cited included the fact that QALYs are not appropriate in all circumstances, that there are issues surrounding the evaluation of multiple comparators, and that genomic testing cannot be addressed in a HTA as if it is a diagnostic test. Several respondents commented that consideration should be given to how personal utility could be captured in economic evaluations. However, there was little consensus about the next steps. One respondent suggested that the first step should be to decide which technologies cannot be appraised using current methods, which non-health outcomes are important, and how these should be captured. Only then should methods change. Another noted that decision-makers should reconsider the evidence that is required to make decisions when they matter (not years afterwards).

6.4.3 Survey conclusions

The results of the survey provide several insights into the objectives and opinions of key stakeholders with respect to the research questions addressed in this thesis. A key finding was that respondents believe that CUAs best capture the costs and benefits of genomic testing in CLL. However, most respondents would prefer to delay an adoption decision with respect to genomic testing until better quality information was available. There was no consensus regarding whether the CBA results were actually helpful to decision-makers in this particular context, although many respondents identified alternative decision contexts in which welfarist approaches would likely be useful. Finally, most respondents believed that standard HTA methods should be adapted to reflect the challenges of conducting

economic evaluations of genetic and genomic tests, although few suggestions were made regarding potential adaptations.

6.5 Conclusions

This chapter has conducted a transparent, objective and thorough comparison of the results of the welfarist and extra-welfarist economic evaluations. Several analytical extensions were undertaken and this facilitated a quantitative comparison between the two sets of results. The adoption decisions suggested by each approach were also examined in detail, and the objectives and opinions of key stakeholders were incorporated into this comparison.

When the results of the different economic evaluation approaches and the stakeholder survey are considered together, it becomes clear that the most likely adoption decision is that genomic testing should not be implemented to guide choice of therapy in this treatment and patient context in CLL. This is an important result. However, the primary research question of this thesis must still be addressed, namely what is the most appropriate economic evaluation approach to use when assessing genomic interventions? A case could be made for both the welfarist and extra-welfarist approaches based on the evidence presented. Chapter Seven therefore concludes this thesis by considering all of this evidence and making some recommendations regarding the most appropriate economic evaluation approach in this context.

Chapter Seven | Conclusion

7.1 Introduction

This chapter begins by presenting an overview of the findings in this thesis before addressing the primary research question: what is the most appropriate economic evaluation approach to use when assessing genomic interventions? The practical implications, generalisability and policy consequences of these findings are then considered, along with the challenges that arose whilst undertaking this research. The chapter ends by highlighting areas for future research.

7.2 Overview of thesis findings

This thesis has explored several issues related to the economic analysis of genomic diagnostic technologies in the UK National Health Service (NHS). The main focus of the thesis was identifying the most appropriate economic evaluation approach to use when assessing genomic interventions. Health economists are currently uncertain as to whether existing extra-welfarist approaches are sufficient to evaluate genomic technologies. Clarification is required because alternative methods, such as the welfarist approach, may lead to different adoption decisions, which will impact on population health.

Two literature reviews were undertaken which described the methodological background to this work. These reviews generated two key findings. First, for every four economic evaluation studies applying both welfarist and extra-welfarist approaches, one shows limited or no concordance between the results of the different approaches. Second, the

quantity and range of methodological challenges in genomics is such that economic evaluations of these technologies present a particular challenge for health economists.

The advantages and disadvantages of the welfarist and extra-welfarist approaches were then demonstrated using a case study: the use of targeted next generation sequencing (NGS) to guide treatment decisions in chronic lymphocytic leukaemia (CLL). There is some evidence that genetic testing could provide clinically relevant information in CLL, and genetic tests can be accessed by CLL specialists in several locations around the UK. However, there is little health economic evidence to support genetic testing in this context. Targeted NGS can provide additional information, over and above that provided by genetic testing, which could be used to better inform CLL treatment decisions. However, the evidence base to support the more widespread use of this technology is currently limited.

The first analysis conducted to address the primary research question was an extra-welfarist economic evaluation of the use of genomic testing in CLL which took two forms: a cost-effectiveness analysis (CEA) and a cost-utility analysis (CUA). The results from both analyses suggest that current practice in the Oxford University Hospitals (OUH) NHS Trust (genetic testing, no ibrutinib treatment) is the most cost-effective strategy at a threshold of £30,000 per life-year/quality-adjusted life year (QALY) gained, hence genomic testing should not be adopted. There is considerable uncertainty surrounding this result, although the strategy with the second highest probability of being cost-effective is also a genetic testing strategy. There is, however, some evidence from the CUA to suggest that genomic testing may be cost-effective for younger patients. Furthermore, the high cost of ibrutinib treatment is likely to be driving these results, rather than the use of genomic testing.

The second analysis conducted to address the primary research question was a welfarist economic evaluation. This took the form of a cost-benefit analysis (CBA), informed by a discrete choice experiment (DCE). This DCE indicated that both clinical and process-related attributes are important to patients, with test effectiveness and speed of result ranked highly. When the results of this DCE were used as inputs into the CBA, the net benefit associated with genomic testing was estimated to be higher than that associated with genetic testing. However, the net benefit for both tests was negative, which complicates the interpretation of these results. A 'what if' benefit transfer analysis was then conducted which indicated that genomic testing should not be adopted in this context.

When the results of the welfarist and extra-welfarist economic evaluations were considered together, it became clear that the most likely adoption decision is that genomic testing should not be implemented to guide treatment decisions in CLL. However, the primary research question must still be addressed. This is tackled in the following section.

7.3 What is the most appropriate economic evaluation approach to use when assessing genomic interventions?

The results of the economic evaluations presented in this thesis had a number of similarities and also several important differences. This section considers these findings in the context of current healthcare decision-making processes in the UK, before making some recommendations regarding the most appropriate approach in a genomic context. An important first step in this process is to consider which factors are important when determining appropriateness.

7.3.1 How can appropriateness be assessed?

There is no single metric which can inform a judgment regarding the suitability of different economic evaluation approaches in specific healthcare contexts. Instead, a decision on the most appropriate approach in genomics must be reached by considering a variety of factors, including the results of case studies relevant to the context in question. These factors include:

- The adoption decisions suggested by each approach in the case study (and whether there is concordance between these decisions);
- Methodological issues related to the outcome measures used in the case study;
- The extent to which these outcome measures capture all of the benefits in this context that are important to patients, or to the health service more generally;
- Whether the economic evaluation approach under consideration is accepted by health technology assessment (HTA) agencies, and allows them to achieve their objectives, as set out by the government;
- Whether the economic evaluation approach under consideration is judged by health economists to be valid in this context;
- Issues surrounding the operationalisation of each approach in this context.

The relative importance of each of these factors will vary for all participants in the healthcare decision-making process, hence there is no objective way to weight these factors when making a recommendation regarding the most appropriate approach in a genomic context. However, some of these factors would appear to be more important than others if a change in the healthcare decision-making process is being considered. Such

factors include whether the approach meets the objectives of HTA agencies and whether there is any evidence that a change in approach will impact on adoption decisions. Hence, although all of the factors in the previous list were considered when formulating a recommendation, these latter two factors were ascribed particular importance.

7.3.2 Which economic evaluation approach is most appropriate in this context?

Table 7.1 summarises the factors in the list in Section 7.3.1 for both the welfarist and extra-welfarist economic evaluations.

Table 7.1: Selecting the most appropriate economic evaluation approach to use when assessing genomic interventions – factors to consider

Factor ^a	Economic evaluation approach	
	Extra-welfarism	Welfarism
Suggested adoption decision in the case study	Do not adopt genomic testing.	Do not adopt genomic testing.
Methodological issues related to the outcome measure used in the case study	<u>LYs</u>: The evidence base underlying the mortality estimates is not fully developed. <u>QALYs</u>: Likely to be considered the most acceptable outcome measure by decision-makers – the utility weights are CLL-specific, although preferences were elicited from the general public, not CLL patients.	Preferences were elicited from CLL patients, were in line with prior expectations, and were similar to those reported by other genetic DCEs. However, there were differences by use of risk graphics, and issues surrounding the levels for the reliability attribute in the DCE.
Does the outcome measure used in the case study capture all of the benefits that are important to <u>patients</u> in this context?	<u>LYs</u>: Somewhat – quality of life may be more important than mortality in a chronic disease context. <u>QALYs</u>: Somewhat – non-health outcomes are also important to patients in this context and are probably not captured by the QALYs.	Somewhat – respondent preferences likely provide information about both health and non-health outcomes. However, health effects arising from treatment were not directly captured.

Factor ^a	Economic evaluation approach	
	Extra-welfarism	Welfarism
Does the outcome measure used in the case study capture all of the benefits that are important to the <u>health service</u> more generally in this context?	<u>LYs</u> : Somewhat – quality of life outcomes also likely to be important to the health service. <u>QALYs</u> : Yes – non-health outcomes are less important for health-maximising decision-makers.	No – some health outcomes are captured, but not those associated with treatment. The non-health outcomes that are captured are less important for health-maximising decision-makers.
Is the approach accepted by HTA agencies?	<u>CEA</u> : Yes, although CUAs are preferred. <u>CUA</u> : Yes, this approach is recommended by NICE.	No, in almost all cases.
Does the approach meet the objectives of HTA agencies, as set out by government?	Yes – CUA allows agencies such as NICE to make recommendations which, when implemented, facilitate the maximisation of health effects.	No – agencies such as NICE are not asked to maximise societal welfare with reference to a societal budget constraint.
Is the economic evaluation approach under consideration judged by health economists to be valid in this context?	Generally yes – in the stakeholder survey most respondents identified CUA as being the most appropriate approach in this context, although most also said that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations in genomics.	Generally no – this approach was criticised in the stakeholder survey for being difficult to interpret. Respondents also noted that methods for determining WTP are not sufficiently reliable.
Are there any pertinent issues surrounding the operationalisation of each approach in this context?	Status quo, so minimises issues surrounding operationalisation. However, some issues were identified in the stakeholder survey (e.g. QALYs may be inappropriate in some circumstances, personal utility is not considered).	Multiple issues were identified in the stakeholder survey (e.g. uncertainty over interpretation and which result is believed, training and capacity issues).

CEA = cost-effectiveness analysis; CLL = chronic lymphocytic leukaemia; CUA = cost-utility analysis; DCE = discrete choice experiment; LYs = life-years; NICE = National Institute of Health and Care Excellence; QALYs = quality-adjusted life years; WTP = willingness-to-pay. ^a Where required, these factors are considered from the perspective of a decision-maker in a UK setting. However, as HTA guidelines in several other developed countries are similar to those in the UK (e.g. Australia, New Zealand, Canada), the comments in this table are likely transferable to these other settings.

Both approaches suggested the same adoption decision in the case study: do not adopt genomic testing. Most survey respondents recommended delaying a decision to allow further evidence to be collected. QALYs would likely be considered the most acceptable outcome measure by decision-makers in the context of genomic testing in CLL, as the evidence base underlying mortality estimates is not fully developed and there are concerns regarding the validity of the preferences elicited in the DCE. However, none of the outcome measures captured all of the benefits that are important to CLL patients: life-years did not include quality of life, QALYs did not capture non-health outcomes, and the CBA did not capture the health effects arising from treatment following testing. The CUA did, however, capture the benefits that are important to the health service more generally in this context. These findings relate to the case study considered in this thesis, but the characteristics of this case study (i.e. limited clinical evidence base, a requirement for willingness-to-pay (WTP) data for testing and treatment, non-health outcomes valued by patients but not health-maximising decision-makers) may be shared by many other genomic interventions, particularly those which inform treatment decisions (e.g. in colorectal cancer [307]). This case study therefore provides useful information which could help to inform a decision regarding the most appropriate economic evaluation approach in this context.

However, it should also be noted that the approach that best captures the costs and benefits of genomic testing is not necessarily the approach that should inform adoption decisions; there are several additional factors to consider. These include the acceptability of each approach (both to decision-makers and health economists) and the objectives of those using the approach in question. Extra-welfarist approaches (generally CUA) are currently recommended by the National Institute for Health and Care Excellence (NICE) in

the UK, whereas welfarist approaches are not generally permitted in HTA submissions (with exceptions made in some public health HTAs). This decision is linked to the objectives of NICE, who are tasked with maximising health effects given the available health budget, not maximising societal welfare. Indeed, Drummond *et al.* note that “*economic evaluation cannot be used to make claims about social welfare or the optimality or otherwise of the resources allocated to health care. Its role is more modest, claiming to inform social decisions in health rather than prescribing social choice*” [308]. The survey results indicated that health economists also believe that extra-welfarist approaches are more valid in this context than welfarist approaches.

Issues surrounding the operationalisation of each approach in a genomic context are an additional consideration. Although health economists did express some methodological concerns with respect to the extra-welfarist approach, this is the status quo and thus it is relatively straightforward to apply in a genomic context. Operationalising welfarism may be more difficult: work is required to strengthen the methodological basis of this approach, and there are training and capacity issues related to carrying out and interpreting these analyses. In addition, permitting the more widespread submission of welfarist economic evaluations to inform HTA decisions presents complications for HTA agencies, requiring them to spend time considering the appropriate evaluative framework at the early stages of HTAs. Decision-making may also be more difficult, particularly when the results of different economic evaluation approaches are not concordant.

Overall, when the evidence in this thesis is considered in the context of current healthcare decision-making processes in the UK, the answer to the primary research question is clear:

the most appropriate economic evaluation approach to use when assessing genomic interventions remains the extra-welfarist approach. The results of the case study analyses do not support the more widespread use of welfarist approaches, there does not appear to be either an appetite or an intellectual argument for change amongst health economists and other stakeholders, and the decision-making context continues to require the use of extra-welfarist approaches. All three of these factors would need to change for welfarist approaches to become a viable option in this context. There is little evidence to suggest that this is likely in the near future (although Drummond *et al.* recently noted that the decision-making context is slowly evolving, and the issue for most decision-makers is no longer 'should we use economic evaluation?', but 'how do we best use it in our setting?' [308]). Furthermore, as noted in the stakeholder survey, if welfarist approaches were considered to be a viable option in a genomic context, there is an argument that such approaches should be considered valid in all clinical contexts. Ultimately, funding for genomic testing and any associated treatment will come from the NHS budget, hence the analytical approach that is applied should be the same as that which is used to evaluate other interventions funded from the same budget.

7.4 Limitations of this research

The analyses presented in this thesis have a number of limitations which should be noted. In terms of the extra-welfarist economic evaluation, the evidence base underlying the transition probabilities in the Markov model is not fully developed, and several simplifying assumptions were required, particularly regarding the estimation of treatment effects by genomic subgroup. These assumptions were, however, tested in sensitivity analyses and the base case results were relatively insensitive to variations in these parameter values. In

addition, the clinical pathways which informed the Markov model and the microcosting studies of genetic and genomic testing were based on current practice at the OUH NHS Trust. As this Trust has teaching responsibilities, it is possible that clinical practice in Oxford is not representative of practice in other Trusts. However, clinicians from other Trusts were asked to comment on the pathways and few differences were identified. Furthermore, no other UK centres currently offer genomic testing in a CLL context.

There were also several limitations related to the DCE. The comparison of different approaches to presenting risk information yielded estimates of WTP for genetic and genomic testing that were very different. A pooled estimate of WTP was used in the CBA, but in the absence of qualitative information on how DCE respondents made their choices it is not clear that this reflects 'true' WTP in this context. A second limitation is that the DCE was undertaken before the underlying clinical study was finalised. This meant that the levels used for some attributes (particularly the reliability attribute) were not well-aligned with the actual values of these attributes. Future studies should undertake such analyses consecutively to avoid this issue, although this is not always possible.

7.5 Implications of this research

This section considers the implications of the research findings reported in this thesis, focusing in particular on the contributions made to the field of health economics, the generalisability of these findings, and the practical and policy consequences of these findings.

7.5.1 Contribution to health economics and other fields

This thesis has made a number of contributions to the health economics literature. First, both of the literature reviews presented in Chapter Two have already been published in peer-reviewed journals [309,310]. The first of these reviews identified only a small number of studies which presented the results of both welfarist and extra-welfarist economic evaluations. This thesis consequently makes a useful contribution to this literature by conducting a CEA, CUA and CBA in the same decision context. There are even fewer examples of CBAs informed by a DCE, hence the welfarist economic evaluation presented in this thesis also contributes to the economic evaluation methodology literature.

Beyond economic evaluation methodology, the finding that estimates of WTP differ when risk presentation is varied is important, and may influence the design of future stated preference studies. The finding that process-related attributes are important to patients in this context has been replicated in other genetic and genomic DCEs, and thus provides further justification for studies which are attempting to incorporate process utility into QALY estimation.

Finally, the clinical evidence base in CLL is currently evolving rapidly, and new treatments such as ibrutinib are being adopted in clinical practice without the support of rigorously conducted economic evaluations. This thesis therefore also provides valuable evidence on the likely cost-effectiveness of these novel high cost therapies.

7.5.2 Generalisability of these findings

The analyses presented in this thesis were based on a case study of the use of targeted NGS in CLL in the UK. It was noted in Section 7.3 that the characteristics of the case study may be shared by other genomic interventions, particularly those which inform treatment decisions. The case study results may therefore be broadly indicative of the likely cost-effectiveness of other genomic tests which inform the stratification of patients to high cost biological therapies, particularly in other chronic disease or cancer settings such as colorectal cancer [307].

More generally, the findings in this thesis are likely to be relevant to health economists and outcomes researchers evaluating other interventions with both health and non-health consequences (e.g. complex interventions). These interventions present a number of similar challenges and are evaluated in the same decision-making context as genomic tests, hence the conclusions reached in this thesis are likely to be of interest.

Finally, these economic evaluations were undertaken in a UK setting. However, the results are likely to be relevant in other countries in which HTA agencies evaluate healthcare interventions using similar methods (e.g. the Canadian Agency for Drugs and Technologies in Health, or the Pharmaceutical Benefits Advisory Committee in Australia).

7.5.3 Practical implications

Chapter Two described a variety of methodological challenges that are faced when conducting economic evaluations in a genomic context. Several of these challenges were considered in this thesis, and these findings have a number of practical implications.

First, the extra-welfarist economic evaluations were conducted from both a NHS and a societal perspective. Although the results were similar in both cases, when younger patients were considered from a societal perspective there was evidence to suggest that genomic testing might be cost-effective. Future studies evaluating genomic tests to guide treatment decisions in older patients from this analytical perspective should therefore also conduct sensitivity analyses in younger patients, as the cost savings from improving productivity and avoiding mortality in this patient group may be noteworthy. In addition, some societal costs that are likely important to older patients were not captured in this analysis, including the costs associated with admission to a nursing home. The use of ibrutinib treatment may prevent such admissions as patients with effectively treated disease are better able to cope with other comorbidities. However, this effect is not captured in the analysis. If it were captured, it is possible that strategies which include ibrutinib treatment would also be cost-effective in older patients when a societal perspective is considered. Future studies evaluating genomic tests to guide treatment decisions in older patients from this analytical perspective would therefore benefit from the inclusion of such costs.

Second, different analytical timeframes were considered in the CEA and CUA. However, these variations only changed the base case analysis results when very short timeframes were evaluated. Future studies should avoid such timeframes unless they are clinically appropriate.

Third, the extra-welfarist economic evaluation considered five alternative test-treatment strategies. However, this was not an exhaustive list of potential strategies. Analysts should

note that only considering a subset of all possible strategies in a genomic economic evaluation adds an additional layer of uncertainty to the study results.

Fourth, Chapter Two identified the need to collect a wide range of cost data in a genomic context. This did not appear to be a particular issue for the case study in this thesis.

Fifth, capturing information on non-health outcomes was expected to be important when evaluating a genomic test. This was borne out by the results of the DCE, which indicated that CLL patients place a high value on process-related attributes. Future studies should be aware of the importance that patients attach to these non-health attributes, even in a scenario in which an intervention yields measurable health benefits.

Finally, it was noted that effectiveness data can be challenging to incorporate into genomic economic evaluations. The data used in this case study did not present a particular problem in this regard. However, these results were not finalised until the underlying clinical study concluded, which reduced the time available to construct the Markov model. Health economists should be aware of this practical issue when planning future studies in a genomic context.

7.5.4 Policy consequences

The findings reported in this thesis have a number of consequences for policymaking. The main conclusion is that the extra-welfarist approach is most appropriate when assessing genomic interventions. This may bring some clarity to policymaking in this arena, although

this conclusion is only informed by a single case study. Further studies evaluating welfarist and extra-welfarist approaches side-by-side in this and other contexts would be helpful.

The results in this thesis also suggest that it would be useful for policymakers to request that data on patient preferences for genomic tests are included in future HTA submissions. Such data may not ultimately be used to inform welfarist economic evaluations, but it will be helpful when designing genomic testing services and may help to ensure that the health benefits of genomic testing are fully realised.

Finally, those who responded to the stakeholder survey made a number of comments which could usefully inform future policymaking in this arena. These results will be made available to policymakers via a future peer-reviewed journal article.

7.6 Future research

A final consideration is the scope for future research which builds on the findings of this thesis. Three areas in which further research would perhaps have the greatest impact are related to widening the evaluative space of the welfarist approach, improving the collection of WTP data and expanding the clinical evidence base in CLL.

First, although this thesis concludes that the extra-welfarist approach is most appropriate when assessing genomic interventions, there is a notable caveat: survey respondents agreed that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations in genomics. Widening the evaluative space of the extra-welfarist approach is one potential methodological variation which may address some of

these concerns, and which may be relatively straightforward to accommodate in current decision-making processes. Options include further developing and operationalising the capability approach (e.g. capability-adjusted life years [311]), or directly incorporating process utility into QALY estimation (although further work is needed in this area [312]).

Second, the results in Chapter Five indicated that different approaches to risk presentation led to large variations in WTP. However, it was difficult to argue that this difference reflected an improvement in risk comprehension without additional information on how respondents made their choices. Such information could be generated by conducting additional analyses alongside future WTP studies, such as qualitative work (e.g. think aloud studies) and behavioural investigations (e.g. using eye-tracking software in DCEs). Furthermore, it may also be helpful to conduct stated preference studies using multiple approaches simultaneously, in order to ‘triangulate’ WTP estimates. Generating WTP data from revealed preference studies would be particularly useful as these studies provide information on real decisions made by individuals in a genomic context. However, it is difficult to conduct such studies at present as there are few markets in which individuals can purchase genomic tests, and few examples of marketed goods that could act as proxies for genomic interventions. One possibility is to study the preferences of individuals for direct-to-consumer genetic tests such as the 23andMe test (www.23andme.com), although as this is offered by a private company, data availability may be limited.

Finally, the value of information analysis presented in Chapter Six indicated that the benefits associated with further research into genomic testing in CLL likely exceed the cost of undertaking this research. This suggests that research which reduces the uncertainty in

the extra-welfarist economic evaluation results would be worthwhile. Studies which generate data on treatment effects in genomic subgroups would be particularly useful.

7.7 Concluding remarks

This thesis has brought together data on the costs and outcomes associated with the use of genomic testing to guide treatment decisions in CLL to explore issues related to the economic evaluation of genomic technologies. The key finding is that the extra-welfarist approach to economic evaluation is most appropriate when assessing genomic interventions. However, an evolution in methods is required to ensure that the results of these analyses provide decision-makers with sufficient information to facilitate the efficient allocation of scarce healthcare resources in this context. These findings will now be disseminated amongst economists, policymakers, oncologists, scientists, patients and other relevant stakeholders with the aim of promoting this evolution.

List of References

1. National Institute for Health and Care Excellence (2006) Familial breast cancer: the classification and care of women at risk of familial breast cancer in primary, secondary and tertiary care
2. Trusheim MR, Berndt ER, Douglas FL (2007) Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. *Nat Rev Drug Discov* 6: 287-293.
3. Towse A, Ossa D, Veenstra D, Carlson J, Garrison L (2013) Understanding the economic value of molecular diagnostic tests: case studies and lessons learned. *J Pers Med* 3: 288-305.
4. Meuwis MA, Fillet M, Chapelle JP, Malaise M, Louis E, et al. (2008) New biomarkers of Crohn's disease: serum biomarkers and development of diagnostic tools. *Expert Rev Mol Diagn* 8: 327-337.
5. Mardis ER (2008) The impact of next-generation sequencing technology on genetics. *Trends Genet* 24: 133-141.
6. U.S. Food and Drug Administration (2013) Drugs@FDA online database.
7. Ioannidis JPA, Khoury MJ (2013) Are randomized trials obsolete or more important than ever in the genomic era? *Genome Medicine* 5: 32.
8. Genomic Health (2012) Annual Report.
9. National Institute for Health and Care Excellence (2013) Gene expression profiling and expanded immunohistochemistry tests for guiding adjuvant chemotherapy decisions in early breast cancer management: MammaPrint, Oncotype DX, IHC4 and Mammostrat.
10. Clyne M, Schully SD, Dotson WD, Douglas MP, Gwinn M, et al. (2014) Horizon scanning for translational genomic research beyond bench to bedside. *Genet Med* 16: 535-538.
11. House of Lords Science and Technology Committee (2009) Genomic Medicine. London.
12. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL (2005) Methods for the Economic Evaluation of Health Care Programmes.
13. National Institute for Health and Care Excellence (2013) Guide to the methods of technology appraisal.
14. Bala MV, Zarkin GA, Mauskopf JA (2002) Conditions for the near equivalence of cost-effectiveness and cost-benefit analyses. *Value in Health* 5: 338-346.
15. Grosse SD, Wordsworth S, Payne K (2008) Economic methods for valuing the outcomes of genetic testing: beyond cost-effectiveness analysis. *Genet Med* 10: 648-654.

16. Gyrd-Hansen D (2005) Willingness to pay for a QALY: Theoretical and methodological issues. *Pharmacoeconomics* 23: 423-432.
17. Johannesson M (1996) Theory and methods of economic evaluation of health care; Johannesson M, editor: *Developments in Health Economics and Public Policy*, vol. 4. Dordrecht; Boston and London: Kluwer Academic.
18. Payne K (2009) Fish and chips all round? Regulation of DNA-based genetic diagnostics. *Health Econ* 18: 1233-1236.
19. van Rooij T, Wilson DM, Marsh S (2012) Personalized medicine policy challenges: measuring clinical utility at point of care. *Expert Rev Pharmacoecon Outcomes Res* 12: 289-295.
20. Cancer Research UK (2012) *CancerStats - Cancer Statistics for the UK*.
21. Catovsky D (2001) *Chronic Lymphocytic Leukaemia*. eLS: John Wiley & Sons, Ltd.
22. Smith A, Howell D, Patmore R, Jack A, Roman E (2011) Incidence of haematological malignancy by sub-type: a report from the Haematological Malignancy Research Network. *Br J Cancer* 105: 1684-1692.
23. Zenz T, Kröber A, Scherer K, Häbe S, Bühler A, et al. (2008) Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up. *Blood* 112: 3322-3329.
24. Knight SJ, Schuh A, Taylor J, Holmes C, Tomlinson I, et al. (2010) Development and application of a targeted array test to diagnose and direct therapy in haematological cancers. *Health Innovation Challenge Fund*.
25. McCormack K, Wake B, Perez J, Fraser C, Cook J, et al. (2005) Laparoscopic surgery for inguinal hernia repair: systematic review of effectiveness and economic evaluation. *Health Technology Assessment* 9: 218.
26. Burr J, Botello Pinzon P, Takwoingi Y, Hernandez R, Vazquez-Montes M, et al. (2012) Surveillance for ocular hypertension: an evidence synthesis and economic evaluation. *Health Technology Assessment* 16: 271.
27. Erdem S, Rigby D (2011) The effects of risk presentations and risk reduction methods on WTP estimates. *International Choice Modelling Conference 2011*.
28. Donaldson C, Gerard K (2005) *Economics of Health Care Financing: The Visible Hand*.
29. Weale A (1998) Rationing health care. *BMJ* 316: 410.
30. www.nhs.uk (2015) Principles and values that guide the NHS.
31. Drummond M, McGuire A (2002) *Economic Evaluation in Health Care: Merging Theory with Practice*: Oxford University Press.

32. Dakin H, Wordsworth S (2013) Cost-minimisation analysis versus cost-effectiveness analysis, revisited. *Health Econ* 22: 22-34.
33. Marshall JD, Harries M, Hill D, Hill CA (2015) PHP139 - Trends in the use of cost-minimization analysis in Economic Assessments submitted to the SMC. *Value in Health* 18: A94.
34. Simoens S (2011) Biosimilar medicines and cost-effectiveness. *ClinicoEconomics and Outcomes Research: CEOR* 3: 29-36.
35. Briggs AH, O'Brien BJ (2001) The death of cost-minimization analysis? *Health Econ* 10: 179-184.
36. Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S (2011) *Applied Methods of Cost-effectiveness Analysis in Health Care*: Oxford University Press.
37. Donaldson C, Birch S, Gafni A (2002) The distribution problem in economic evaluation: income and the valuation of costs and consequences of health care programmes. *Health Econ* 11: 55-70.
38. Blumenschein K, Johannesson M (1996) Economic evaluation in healthcare. A brief history and future directions. *Pharmacoeconomics* 10: 114-122.
39. Johannesson M (1995) The Relationship between Cost-Effectiveness Analysis and Cost-Benefit-Analysis. *Social Science & Medicine* 41: 483-489.
40. Evans KW (1997) Economic evaluation of oral sumatriptan compared with oral caffeine/ergotamine for migraine. *Pharmacoeconomics* 12: 565-577.
41. Johannesson M, Jonsson B (1991) Economic evaluation in health care: Is there a role for cost-benefit analysis? *Health Policy* 17: 1-23.
42. Williams A (1985) Economics of coronary artery bypass grafting. *Br Med J (Clin Res Ed)* 291: 326-329.
43. Gafni A (2006) Economic Evaluation of Health-care Programmes: Is CEA Better than CBA? *Environmental & Resource Economics* 34: 407-418.
44. Brent RJ (2003) *Cost-benefit analysis and health care evaluations*; Brent RJ, editor: Cheltenham, U.K. and Northampton, Mass.: Elgar; distributed by American International Distribution Corporation, Williston, Vt.
45. Tsuchiya A, Williams A (2001) Welfare economic and economic evaluation. In: Drummond M, McGuire A, editors. *Economic evaluation in health care: merging theory with practice*: Oxford University Press.
46. Raftery J (1998) Economic evaluation: an introduction. *BMJ* 316: 1013-1014.
47. Brouwer WB, Koopmanschap MA (2000) On the economic foundations of CEA. Ladies and gentlemen, take your positions! *J Health Econ* 19: 439-459.

48. Smith R, Lorgelly P, Al-Janabi H, Venkatapuram S, Coast J (2012) The Capability Approach: An Alternative Evaluation Paradigm for Health Economics? In: Jones AM, editor. *The Elgar Companion to Health Economics*. Second ed: Edward Elgar.
49. Smith RD, Sach TH (2010) Contingent valuation: what needs to be done? *Health Econ Policy Law* 5: 91-111.
50. Bryan S, Dolan P (2004) Discrete choice experiments in health economics. For better or for worse? *Eur J Health Econ* 5: 199-202.
51. Baker R, Currie GR, Donaldson C (2010) What needs to be done in contingent valuation: have Smith and Sach missed the boat? *Health Econ Policy Law* 5: 113-121.
52. Mooney G (1998) Beyond health outcomes: The benefits of health care. *Health Care Analysis* 6: 99-105.
53. Mishan E (1971) *Cost-benefit analysis*. Sydney: Allen and Unwin.
54. Brent RJ (2002) A Simple Method for Converting a Cost-Effectiveness Analysis to a Cost-Benefit Analysis with an Application to State Mental Health Expenditures. *Public Finance Review* 30: 144-160.
55. Conrad DA, Deyo RA (1994) Economic decision analysis in the diagnosis and treatment of low back pain. A methodologic primer. *Spine* 19: 2101S-2106S.
56. Kenkel D (2006) WTP- and QALY-based approaches to valuing health for policy: Common ground and disputed territory. *Environmental & Resource Economics* 34: 419-437.
57. Phelps CE, Mushlin AI (1991) On the (near) equivalence of cost-effectiveness and cost-benefit analyses. *Int J Technol Assess Health Care* 7: 12-21.
58. Donaldson C (1998) The (near) equivalence of cost-effectiveness and cost-benefit analyses. Fact or fallacy? *Pharmacoeconomics* 13: 389-396.
59. Dolan P, Edlin R (2002) Is it really possible to build a bridge between cost-benefit analysis and cost-effectiveness analysis? *J Health Econ* 21: 827-843.
60. Bleichrodt H, Quiggin J (1999) Life-cycle preferences over consumption and health: When is cost-effectiveness analysis equivalent to cost-benefit analysis? *J Health Econ* 18: 681-708.
61. Olsen JA, Donaldson C (1998) Helicopters, hearts and hips: using willingness to pay to set priorities for public sector health care programmes. *Soc Sci Med* 46: 1-12.
62. Payne K, McAllister M, Davies LM (2013) Valuing the economic benefits of complex interventions: when maximising health is not sufficient. *Health Econ* 22: 258-271.
63. Medical Research Council. (2008) *Developing and evaluating complex interventions: new guidance*. London: MRC.

64. Hardeman W, Sutton S, Griffin S, Johnston M, White A, et al. (2005) A causal modelling approach to the development of theory-based behaviour change programmes for trial evaluation. *Health Education Research* 20: 676-687.
65. Garau M, Towse A, Garrison L, Housman L, Ossa D (2012) Can and should value based pricing be applied to molecular diagnostics? OHE research paper 12/03.
66. The Academy of Medical Sciences (2013) Realising the potential of stratified medicine.
67. National Institute for Health and Care Excellence (2012) Methods for the development of NICE public health guidance (third edition).
68. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, et al. (2013) Consolidated Health Economic Evaluation Reporting Standards (CHEERS) Statement. *Pharmacoeconomics* 31: 361-367.
69. Abdul Pari AA, Simon J, Wolstenholme J, Geddes JR, Goodwin GM (2014) Economic evaluations in bipolar disorder: a systematic review and critical appraisal. *Bipolar Disorders* 16: 557-582.
70. Beale SJ, Bending MW, Trueman P, Naidoo B (2012) Should we invest in environmental interventions to encourage physical activity in England? An economic appraisal. *Eur J Public Health* 22: 869-873.
71. Brisson M, Edmunds WJ (2006) Impact of model, methodological, and parameter uncertainty in the economic analysis of vaccination programs. *Med Decis Making* 26: 434-446.
72. Chen S, Shaheen A, Garber A (1998) Cost-effectiveness and cost-benefit analysis of using methotrexate vs Goeckerman therapy for psoriasis: A pilot study. *Archives of Dermatology* 134: 1602-1608.
73. Yianni J, Green AL, McIntosh E, Bittar RG, Joint C, et al. (2005) The costs and benefits of deep brain stimulation surgery for patients with dystonia: An initial exploration. *Neuromodulation* 8: 155-161.
74. Torrance G, Walker V, Grossman R, Mukherjee J, Vaughan D, et al. (1999) Economic evaluation of ciprofloxacin compared with usual antibacterial care for the treatment of acute exacerbations of chronic bronchitis in patients followed for 1 year. *Pharmacoeconomics* 16: 499-520.
75. Holtgrave DR, Valdiserri RO, Gerber AR, Hinman AR (1993) Human immunodeficiency virus counseling, testing, referral, and partner notification services. A cost-benefit analysis. *Arch Intern Med* 153: 1225-1230.
76. Loisel P, Lemaire J, Poitras S, Durand MJ, Champagne F, et al. (2002) Cost-benefit and cost-effectiveness analysis of a disability prevention model for back pain management: a six year follow up study. *Occup Environ Med* 59: 807-815.

77. McKay NL, Phillips KM (1991) An economic evaluation of mandatory premarital testing for HIV. *Inquiry* 28: 236-248.
78. Moradi-Lakeh M, Shakerian S, Esteghamati A (2012) Immunization against *Haemophilus Influenzae* Type b in Iran; Cost-utility and Cost-benefit Analyses. *Int J Prev Med* 3: 332-340.
79. Scott RD, 2nd, Meltzer MI, Erickson LJ, De Wals P, Rosenstein NE (2002) Vaccinating first-year college students living in dormitories for Meningococcal disease: an economic analysis. *Am J Prev Med* 23: 98-105.
80. van Oostrom SH, Heymans MW, de Vet HC, van Tulder MW, van Mechelen W, et al. (2010) Economic evaluation of a workplace intervention for sick-listed employees with distress. *Occup Environ Med* 67: 603-610.
81. Driessen M, Bosmans J, Proper K, Anema J, Bongers P, et al. (2012) The economic evaluation of a Participatory Ergonomics programme to prevent low back and neck pain. *Work* 41 Suppl 1: 2315-2320.
82. Johannesson M, Fagerberg B (1992) A health-economic comparison of diet and drug treatment in obese men with mild hypertension. *Journal of Hypertension* 10: 1063-1070.
83. Liu Y, Mack KA, Diekman ST (2012) Smoke alarm giveaway and installation programs: an economic evaluation. *Am J Prev Med* 43: 385-391.
84. Zarnke KB, Levine MA, O'Brien BJ (1997) Cost-benefit analyses in the health-care literature: don't judge a study by its label. *J Clin Epidemiol* 50: 813-822.
85. Cohen DR, Patel N (2009) The potential to forgo social welfare gains through overreliance on cost effectiveness/cost utility analyses in the evidence base for public health. *J Environ Public Health* 2009: 107927.
86. McIntosh E, Clarke PM, Frew E, Louviere JL, editors (2010) *Applied Methods of Cost-benefit Analysis in Health Care*: Oxford University Press.
87. Phillips KA, Veenstra D, Van Bebbber S, Sakowski J (2003) An introduction to cost-effectiveness and cost-benefit analysis of pharmacogenomics. *Pharmacogenomics* 4: 231-239.
88. PHG Foundation (2011) *Next Steps In The Sequence: The Implications Of Whole Genome Sequencing For Health In The UK*.
89. United States Centers for Disease Control and Prevention (2012) *Genomic Testing*.
90. Grosse SD, Rogowski WH, Ross LF, Cornel MC, Dondorp WJ, et al. (2010) Population Screening for Genetic Disorders in the 21st Century: Evidence, Economics, and Ethics. *Public Health Genomics* 13: 106-115.
91. Rogowski WH, Grosse SD, John J, Kaariainen H, Kent A, et al. (2010) Points to consider in assessing and appraising predictive genetic tests. *J Community Genet* 1: 185-194.

92. Krol M, Brouwer W, Rutten F (2013) Productivity costs in economic evaluations: past, present, future. *Pharmacoeconomics* 31: 537-549.
93. Giacomini M, Miller F, O'Brien BJ (2003) Economic considerations for health insurance coverage of emerging genetic tests. *Community Genetics* 6: 61-73.
94. Djalalov S, Musa Z, Mendelson M, Siminovitch K, Hoch J (2011) A review of economic evaluations of genetic testing services and interventions (2004-2009). *Genet Med* 13: 89-94.
95. Evans JP, Berg JS, Olshan AF, Magnuson T, Rimer BK (2013) We screen newborns, don't we?: realizing the promise of public health genomics. *Genet Med* 15: 332-334.
96. Rogowski W (2007) Current impact of gene technology on healthcare. A map of economic assessments. *Health Policy* 80: 340-357.
97. Rogowski WH (2009) The cost-effectiveness of screening for hereditary hemochromatosis in Germany: a remodeling study. *Med Decis Making* 29: 224-238.
98. Rogowski WH, Grosse SD, Khoury MJ (2009) Challenges of translating genetic tests into clinical and public health practice. *Nat Rev Genet* 10: 489-495.
99. Goddard KAB, Knaus WA, Whitlock E, Lyman GH, Feigelson HS, et al. (2012) Building the evidence base for decision making in cancer genomic medicine using comparative effectiveness research. *Genetics in Medicine* 14: 633-642.
100. Lin X, Tang W, Ahmad S, Lu J, Colby CC, et al. (2012) Applications of targeted gene capture and next-generation sequencing technologies in studies of human deafness and other genetic disabilities. *Hearing Research* 288: 67-76.
101. Conti R, Veenstra DL, Armstrong K, Lesko LJ, Grosse SD (2010) Personalized medicine and genomics: challenges and opportunities in assessing effectiveness, cost-effectiveness, and future research priorities. *Med Decis Making* 30: 328-340.
102. Wetterstrand KA (2013) DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). Available at: www.genome.gov/sequencingcosts.
103. Ransohoff DF, Khoury MJ (2010) Personal genomics: information can be harmful. *Eur J Clin Invest* 40: 64-68.
104. Mihaescu R, van Hoek M, Sijbrands EJ, Uitterlinden AG, Witteman JC, et al. (2009) Evaluation of risk prediction updates from commercial genome-wide scans. *Genet Med* 11: 588-594.
105. Rogowski W (2006) Genetic screening by DNA technology: a systematic review of health economic evidence. *Int J Technol Assess Health Care* 22: 327-337.
106. Roth JA, Garrison LP, Burke W, Ramsey SD, Carlson R, et al. (2011) Stakeholder Perspectives on a Risk-Benefit Framework for Genetic Testing. *Public Health Genomics* 14: 59-67.

107. Chapman AM, Taylor CA, Girling AJ (2014) Early HTA to Inform Medical Device Development Decisions - The Headroom Method. In: Roa Romero LM, editor. XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013: Springer International Publishing. pp. 1151-1154.
108. Snyder SR, Mitropoulou C, Patrinos GP, Williams MS (2014) Economic Evaluation of Pharmacogenomics: A Value-Based Approach to Pragmatic Decision Making in the Face of Complexity. *Public Health Genomics* 17: 256-264.
109. Grosse SD, Kalman L, Khoury MJ (2010) Evaluation of the validity and utility of genetic testing for rare diseases. *Advances in Experimental Medicine and Biology* 686.
110. Alberts SR, Yu T, Behrens RJ, Renfro LA, Srivastava G, et al. (2013) Real-world comparative economics of a 12-gene assay for prognosis in stage II colon cancer. *Journal of Clinical Oncology* 31: 391.
111. Holt S, Bertelli G, Humphreys I, Valentine W, Durrani S, et al. (2013) A decision impact, decision conflict and economic assessment of routine Oncotype DX testing of 146 women with node-negative or pN1mi, ER-positive breast cancer in the U.K. *Br J Cancer* 108: 2250-2258.
112. Wordsworth S, Buchanan J, Papanicolas I, Taylor J, Frayling I, et al. (2008) Molecular testing for somatic mutations in common cancers: the views of UK oncologists. *Journal of Clinical Pathology* 61: 761-765.
113. Hall PS, McCabe C, Stein RC, Cameron D (2012) Economic evaluation of genomic test-directed chemotherapy for early-stage lymph node-positive breast cancer. *J Natl Cancer Inst* 104: 56-66.
114. Sanderson S, Zimmern R, Kroese M, Higgins J, Patch C, et al. (2005) How can the evaluation of genetic tests be enhanced? Lessons learned from the ACCE framework and evaluating genetic tests in the United Kingdom. *Genet Med* 7: 495-500.
115. Ladabaum U, Wang G, Terdiman J, Blanco A, Kuppermann M, et al. (2011) Strategies to identify the Lynch syndrome among patients with colorectal cancer. *Annals of Internal Medicine*. pp. 69-79.
116. Mvundura M, Grosse SD, Hampel H, Palomaki GE (2010) The cost-effectiveness of genetic testing strategies for Lynch syndrome among newly diagnosed patients with colorectal cancer. *Genet Med* 12: 93-104.
117. McCormick P. Development of personalised medicines for rare genetic conditions: some problems; 2012; London, UK.
118. Feero W, Wicklund C, Veenstra DL (2013) The economics of genomic medicine: Insights from the iom roundtable on translating genomic-based research for health. *JAMA* 309: 1235-1236.

119. Payne K (2007) Towards an economic evidence base for pharmacogenetics: consideration of outcomes is key. *Pharmacogenomics* 9: 1-4.
120. Faulkner E, Annemans L, Garrison L, Helfand M, Holtorf A-P, et al. (2012) Challenges in the Development and Reimbursement of Personalized Medicine—Payer and Manufacturer Perspectives and Implications for Health Economics and Outcomes Research: A Report of the ISPOR Personalized Medicine Special Interest Group. *Value in Health* 15: 1162-1171.
121. Singer DRJ, Watkins J (2012) Using companion and coupled diagnostics within strategy to personalize targeted medicines. *Personalized Medicine* 9: 751-761.
122. Veenstra DL, Higashi MK, Phillips KA (2000) Assessing the cost-effectiveness of pharmacogenomics. *AAPS PharmSci* 2: E29.
123. Wordsworth S, Buchanan J, Regan R, Davison V, Smith K, et al. (2007) Diagnosing idiopathic learning disability: a cost-effectiveness analysis of microarray technology in the National Health Service of the United Kingdom. *Genomic Med* 1: 35-45.
124. Vataire AL, Laas E, Aballea S, Gligorov J, Rouzier R, et al. (2012) [Cost-effectiveness of a chemotherapy predictive test]. *Bull Cancer* 99: 907-914.
125. Deverka PA, Schully SD, Ishibe N, Carlson JJ, Freedman A, et al. (2012) Stakeholder assessment of the evidence for cancer genomic tests: insights from three case studies. *Genet Med* 14: 656-662.
126. Malhotra AK, Zhang JP, Lencz T (2012) Pharmacogenetics in psychiatry: Translating research into clinical practice. *Molecular Psychiatry* 17: 760-769.
127. Van Den Akker-Van Marle ME, Gurwitz D, Detmar SB, Enzing CM, Hopkins MM, et al. (2006) Cost-effectiveness of pharmacogenomics in clinical practice: A case study of thiopurine methyltransferase genotyping in acute lymphoblastic leukemia in Europe. *Pharmacogenomics* 7: 783-792.
128. De Leon J (2009) Pharmacogenomics: The promise of personalized medicine for CNS disorders. *Neuropsychopharmacology* 34: 159-172.
129. Knight SJ, Yau C, Clifford R, Timbs AT, Sadighi Akha E, et al. (2012) Quantification of subclonal distributions of recurrent genomic aberrations in paired pre-treatment and relapse samples from patients with B-cell chronic lymphocytic leukemia. *Leukemia* 26: 1564-1575.
130. Gurwitz D, Zika E, Hopkins MM, Gaisser S, Ibarreta D (2009) Pharmacogenetics in Europe: barriers and opportunities. *Public Health Genomics* 12: 134-141.
131. Bajaj PS, Veenstra DL (2013) A risk-benefit analysis of factor V Leiden testing to improve pregnancy outcomes: a case study of the capabilities of decision modeling in genomics. *Genet Med* 15: 374-381.

132. Jarrett J, Mugford M (2006) Genetic health technology and economic evaluation: a critical review. *Appl Health Econ Health Policy* 5: 27-35.
133. Veenstra DL, Piper M, Haddow JE, Pauker SG, Klein R, et al. (2013) Improving the efficiency and relevance of evidence-based recommendations in the era of whole-genome sequencing: an EGAPP methods update. *Genet Med* 15: 14-24.
134. Guzauskas GF, Garrison LP, Stock J, Au S, Doyle DL, et al. (2013) Stakeholder perspectives on decision-analytic modeling frameworks to assess genetic services policy. *Genet Med* 15: 84-87.
135. Foster MW, Mulvihill JJ, Sharp RR (2009) Evaluating the utility of personal genomic information. *Genet Med* 11: 570-574.
136. Grosse SD, Khoury MJ (2006) What is the clinical utility of genetic testing? *Genet Med* 8: 448-450.
137. Asch DA, Hershey JC, Pauly MV, Patton JP, Jedrzejewski MK, et al. (1996) Genetic screening for reproductive planning: methodological and conceptual issues in policy analysis. *Am J Public Health* 86: 684-690.
138. Caughey AB (2005) Cost-effectiveness analysis of prenatal diagnosis: methodological issues and concerns. *Gynecol Obstet Invest* 60: 11-18.
139. Grosse SD, McBride CM, Evans JP, Khoury MJ (2009) Personal utility and genomic information: look before you leap. *Genet Med* 11: 575-576.
140. Khoury MJ, Gwinn M, Dotson WD, Bowen MS (2011) Is there a need for PGxceptionalism? *Genetics in Medicine* 13: 866-867.
141. Flynn TN (2010) Valuing citizen and patient preferences in health: recent developments in three types of best-worst scaling. *Expert Rev Pharmacoecon Outcomes Res* 10: 259-267.
142. Najafzadeh M, Johnston KM, Peacock SJ, Connors JM, Marra MA, et al. (2013) Genomic testing to determine drug response: measuring preferences of the public and patients using Discrete Choice Experiment (DCE). *BMC Health Serv Res* 13: 454.
143. Food and Drug Administration (2015) Patient Preference Information – Submission, Review in PMAs, HDE Applications, and De Novo Requests, and Inclusion in Device Labeling: Draft Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders.
144. Rogowski W, Payne K, Schnell-Inderst P, Manca A, Rochau U, et al. (2015) Concepts of 'personalization' in personalized medicine: implications for economic evaluation. *Pharmacoeconomics* 33: 49-59.

145. Issa AM (2008) Evaluating the value of genomic diagnostics: Implications for clinical practice and public policy. *Advances in Health Economics and Health Services Research* 19.
146. Issa A, Tufail W, Ajike R, Tenorio J, McKeever J (2011) Breast and colorectal cancer patients' preferences for pharmacogenomics: A discrete choice experiment. 27th International Conference on Pharmacoepidemiology and Therapeutic Risk Management. Chicago, IL, United States.
147. Van Bebber SL, Liang S-Y, Phillips KA, Marshall D, Walsh J, et al. (2007) Valuing personalized medicine: willingness to pay for genetic testing for colorectal cancer risk. *Personalized Medicine* 4: 341-350.
148. Phillips KA, Liang SY, Van Bebber S, Grp CR (2008) Challenges to the translation of genomic information into clinical practice and health policy: Utilization, preferences and economic value. *Current Opinion in Molecular Therapeutics* 10: 260-266.
149. Thariani R, Veenstra DL, Carlson JJ, Garrison LP, Ramsey S (2012) Paying for personalized care: Cancer biomarkers and comparative effectiveness. *Molecular Oncology* 6: 260-266.
150. Jacob I, Awada AH, Payne K, Annemans L (2013) Stratified medicine: a call for action. *Expert Rev Pharmacoecon Outcomes Res* 13: 277-279.
151. Khoury MJ, Rich EC, Randhawa G, Teutsch SM, Niederhuber J (2009) Comparative effectiveness research and genomic medicine: an evolving partnership for 21st century medicine. *Genet Med* 11: 707-711.
152. Beauchamp JP, Cesarini D, Johannesson M, van der Loos MJ, Koellinger PD, et al. (2011) Molecular Genetics and Economics. *J Econ Perspect* 25: 57-82.
153. Rogowski WH, Grosse SD, Meyer E, John J, Palmer S (2012) Using value of information analysis in decision making about applied research : The case of genetic screening for hemochromatosis in Germany. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz* 55: 700-709.
154. Flowers CR, Veenstra D (2004) The role of cost-effectiveness analysis in the era of pharmacogenomics. *Pharmacoeconomics* 22: 481-493.
155. Siebert U, Taylor R, Schnell-Inderst P, Sculpher M, Iglesias C, et al. (2015) Do we need new methods for comparative effectiveness research of medical devices? *International Health Economics Association 11th World Congress*. Milan.
156. Husereau D, Marshall DA, Levy AR, Peacock S, Hoch JS (2014) Health technology assessment and personalized medicine: are economic evaluation guidelines sufficient to support decision making? *Int J Technol Assess Health Care* 30: 179-187.

157. Shabaruddin FH, Fleeman ND, Payne K (2015) Economic evaluations of personalized medicine: existing challenges and current developments. *Pharmgenomics Pers Med* 8: 115-126.
158. Oscier D, Dearden C, Erem E, Fegan C, Follows G, et al. (2012) Guidelines on the diagnosis, investigation and management of Chronic Lymphocytic Leukaemia. British Committee for Standards in Haematology.
159. Eichhorst B, Hallek M, Dreyling M, Group EGW (2010) Chronic lymphocytic leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 21 Suppl 5: v162-164.
160. Sant M, Allemani C, Tereanu C, De Angelis R, Capocaccia R, et al. (2010) Incidence of hematologic malignancies in Europe by morphologic subtype: results of the HAEMACARE project. *Blood* 116: 3724-3734.
161. Redaelli A, Laskin BL, Stephens JM, Botteman MF, Pashos CL (2004) The clinical and epidemiological burden of chronic lymphocytic leukaemia. *Eur J Cancer Care (Engl)* 13: 279-287.
162. Andritsos L, Khoury H (2002) Chronic lymphocytic leukemia. *Curr Treat Options Oncol* 3: 225-231.
163. Oscier D, Fegan C, Hillmen P, Illidge T, Johnson S, et al. (2004) Guidelines on the diagnosis and management of chronic lymphocytic leukaemia. *British Journal of Haematology* 125: 294-317.
164. Gunz FW, Gunz JP, Veale AM, Chapman CJ, Houston IB (1975) Familial leukaemia: a study of 909 families. *Scand J Haematol* 15: 117-131.
165. Leukaemia & Lymphoma Research (2011) Chronic Lymphocytic Leukaemia (CLL).
166. Brown JR (2008) Inherited predisposition to chronic lymphocytic leukemia. *Expert Rev Hematol* 1: 51-61.
167. Goldin LR, Pfeiffer RM, Li X, Hemminki K (2004) Familial risk of lymphoproliferative tumors in families of patients with chronic lymphocytic leukemia: results from the Swedish Family-Cancer Database. *Blood* 104: 1850-1854.
168. Goldin LR, Slager SL, Caporaso NE (2010) Familial chronic lymphocytic leukemia. *Curr Opin Hematol* 17: 350-355.
169. Houlston RS, Sellick G, Yuille M, Matutes E, Catovsky D (2003) Causation of chronic lymphocytic leukemia--insights from familial disease. *Leuk Res* 27: 871-876.
170. Auer RL, Dighiero G, Goldin LR, Syndercombe-Court D, Jones C, et al. (2007) Trinucleotide repeat dynamic mutation identifying susceptibility in familial and sporadic chronic lymphocytic leukaemia. *Br J Haematol* 136: 73-79.

171. Goldin LR, Sgambati M, Marti GE, Fontaine L, Ishibe N, et al. (1999) Anticipation in familial chronic lymphocytic leukemia. *Am J Hum Genet* 65: 265-269.
172. Horwitz M, Goode EL, Jarvik GP (1996) Anticipation in familial leukemia. *Am J Hum Genet* 59: 990-998.
173. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, et al. (2008) Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood* 111: 5446-5456.
174. Pettitt AR (2006) Making sense of prognostic factors in CLL. CLL Support Association.
175. Patel A, Kang SH, Lennon PA, Li YF, Rao PN, et al. (2008) Validation of a targeted DNA microarray for the clinical evaluation of recurrent abnormalities in chronic lymphocytic leukemia. *Am J Hematol* 83: 540-546.
176. Zenz T, Mertens D, Döhner H, Stilgenbauer S (2011) Importance of genetics in chronic lymphocytic leukemia. *Blood Reviews* 25: 131-137.
177. Gunn SR, Mohammed MS, Gorre ME, Cotter PD, Kim J, et al. (2008) Whole-genome scanning by array comparative genomic hybridization as a clinical tool for risk assessment in chronic lymphocytic leukemia. *J Mol Diagn* 10: 442-451.
178. O'Malley DP, Giudice C, Chang AS, Chang D, Barry TS, et al. (2011) Comparison of array comparative genomic hybridization (aCGH) to FISH and cytogenetics in prognostic evaluation of chronic lymphocytic leukemia. *Int J Lab Hematol* 33: 238-244.
179. Tyybakinoja A, Vilpo J, Knuutila S (2007) High-resolution oligonucleotide array-CGH pinpoints genes involved in cryptic losses in chronic lymphocytic leukemia. *Cytogenet Genome Res* 118: 8-12.
180. Zenz T, Mertens D, Dohner H, Stilgenbauer S (2008) Molecular diagnostics in chronic lymphocytic leukemia - pathogenetic and clinical implications. *Leuk Lymphoma* 49: 864-873.
181. Dohner H, Stilgenbauer S, Benner A, Leupolt E, Krober A, et al. (2000) Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med* 343: 1910-1916.
182. Ouillet P, Collins R, Shakhan S, Li J, Li C, et al. (2011) The prognostic significance of various 13q14 deletions in chronic lymphocytic leukemia. *Clin Cancer Res* 17: 6778-6790.
183. Cuneo A, Rigolin GM, Bigoni R, Angeli CD, Veronese A, et al. (2003) Chronic lymphocytic leukemia with 6q- shows distinct hematological features and intermediate prognosis. *Leukemia* 18: 476-483.
184. Hanrahan AJ, Schultz N, Westfal ML, Sakr RA, Giri DD, et al. (2012) Genomic complexity and AKT dependence in serous ovarian cancer. *Cancer Discov* 2: 56-67.

185. Berger MF, Lawrence MS, Demichelis F, Drier Y, Cibulskis K, et al. (2011) The genomic complexity of primary human prostate cancer. *Nature* 470: 214-220.
186. Zhang L, Znoyko I, Costa LJ, Conlin LK, Daber RD, et al. (2011) Clonal diversity analysis using SNP microarray: a new prognostic tool for chronic lymphocytic leukemia. *Cancer Genetics* 204: 654-665.
187. Pospisilova S, Gonzalez D, Malcikova J, Trbusek M, Rossi D, et al. (2012) ERIC recommendations on TP53 mutation analysis in chronic lymphocytic leukemia. *Leukemia* 26: 1458-1461.
188. Bienz N, Cardy DL, Leyland MJ, Hulten MA (1993) Trisomy 12 in B-cell chronic lymphocytic leukaemia: an evaluation of 33 patients by direct fluorescence in situ hybridization (FISH). *Br J Haematol* 85: 819-822.
189. Coignet L, Bertheas MF, Vasselon C, Jaubert J, Reynaud J, et al. (1993) Trisomy 12 in B-cell chronic lymphocytic leukemia: interphase study by in situ hybridization in 75 patients. *Nouv Rev Fr Hematol* 35: 129-134.
190. Escudier SM, Pereira-Leahy JM, Drach JW, Weier HU, Goodacre AM, et al. (1993) Fluorescent in situ hybridization and cytogenetic studies of trisomy 12 in chronic lymphocytic leukemia. *Blood* 81: 2702-2707.
191. Mason JM, Drummond MF, Bosanquet AG, Sheldon TA (2000) The DiSC ASSAY. *International Journal of Technology Assessment in Health Care* 15: 173-184.
192. Al Zaabi EA, Fernandez LA, Sadek IA, Riddell DC, Greer WL (2010) Multiplex ligation-dependent probe amplification versus multiprobe fluorescence in situ hybridization to detect genomic aberrations in chronic lymphocytic leukemia: a tertiary center experience. *J Mol Diagn* 12: 197-203.
193. Sargent R, Jones D, Abruzzo LV, Yao H, Bonderover J, et al. (2009) Customized oligonucleotide array-based comparative genomic hybridization as a clinical assay for genomic profiling of chronic lymphocytic leukemia. *J Mol Diagn* 11: 25-34.
194. LimeSurvey Project Team / Carsten Schmitz (2015) LimeSurvey: An Open Source survey tool. LimeSurvey Project Hamburg, Germany.
195. Quesada V, Conde L, Villamor N, Ordonez GR, Jares P, et al. (2012) Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet* 44: 47-52.
196. Catovsky D, Richards S, Matutes E, Oscier D, Dyer MJ, et al. (2007) Assessment of fludarabine plus cyclophosphamide for patients with chronic lymphocytic leukaemia (the LRF CLL4 Trial): a randomised controlled trial. *Lancet* 370: 230-239.
197. National Institute for Health and Care Excellence (2010) Ofatumumab for the treatment of chronic lymphocytic leukaemia in patients who are refractory to fludarabine and alemtuzumab.

198. NICE technology appraisal guidance 174 (2009) Rituximab for the first-line treatment of chronic lymphocytic leukaemia.
199. NICE technology appraisal guidance 193 (2010) Rituximab for the treatment of relapsed or refractory chronic lymphocytic leukaemia.
200. Badoux XC, Keating MJ, Wierda WG (2011) What is the best frontline therapy for patients with CLL and 17p deletion? *Curr Hematol Malig Rep* 6: 36-46.
201. Byrd JC, Furman RR, Coutre SE, Flinn IW, Burger JA, et al. (2013) Targeting BTK with ibrutinib in relapsed chronic lymphocytic leukemia. *N Engl J Med* 369: 32-42.
202. Furman RR, Sharman JP, Coutre SE, Cheson BD, Pagel JM, et al. (2014) Idelalisib and rituximab in relapsed chronic lymphocytic leukemia. *N Engl J Med* 370: 997-1007.
203. Souers AJ, Levenson JD, Boghaert ER, Ackler SL, Catron ND, et al. (2013) ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat Med* 19: 202-208.
204. NHS England (2015) National Cancer Drugs Fund List, Version 4.0.
205. Ghia P (2012) A look into the future: can minimal residual disease guide therapy and predict prognosis in chronic lymphocytic leukemia? *ASH Education Program Book* 2012: 97-104.
206. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74: 5463-5467.
207. Voelkerding KV, Dames SA, Durtschi JD (2009) Next-Generation Sequencing: From Basic Research to Diagnostics. *Clin Chem* 55: 641-658.
208. Sikkema-Raddatz B, Johansson LF, de Boer EN, Almomani R, Boven LG, et al. (2013) Targeted Next-Generation Sequencing can Replace Sanger Sequencing in Clinical Diagnostics. *Human Mutation* 34: 1035-1042.
209. Alsolami R, Knight SJL, Schuh A (2013) Clinical application of targeted and genome-wide technologies: can we predict treatment responses in chronic lymphocytic leukemia? *Personalized Medicine* 10: 361-376.
210. Boon EMJ, Faas BHW (2013) Benefits and limitations of whole genome versus targeted approaches for noninvasive prenatal testing for fetal aneuploidies. *Prenatal Diagnosis* 33: 563-568.
211. Mansouri L, Sutton LA, Ljungstrom V, Sorqvist EF, Gunnarsson R, et al. (2014) Feasibility of targeted next-generation sequencing of the TP53 and ATM genes in chronic lymphocytic leukemia. *Leukemia* 28: 694-696.
212. Jethwa A, Hülle J, Stolz T, Blume C, Sellner L, et al. (2013) Targeted resequencing for analysis of clonal composition of recurrent gene mutations in chronic lymphocytic leukaemia. *British Journal of Haematology* 163: 496-500.

213. Improgo MR, Brown JR (2013) Genomic approaches to chronic lymphocytic leukemia. *Hematol Oncol Clin North Am* 27: 157-171.
214. Villamor N, Lopez-Guillermo A, Lopez-Otin C, Campo E (2013) Next-generation sequencing in chronic lymphocytic leukemia. *Semin Hematol* 50: 286-295.
215. Public Health Genetics (2006) Evaluation of the use of array comparative genomic hybridisation in the diagnosis of learning disability: Report of a UK Genetic Testing Network Working Party.
216. Barton P, Bryan S, Robinson S (2004) Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy* 9: 110-118.
217. Brennan A, Chick SE, Davies R (2006) A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 15: 1295-1310.
218. Cooper NJ, Sutton AJ, Abrams KR, Turner D, Wailoo A (2004) Comprehensive decision analytical modelling in economic evaluation: a Bayesian approach. *Health Econ* 13: 203-226.
219. Byrd JC, Brown JR, O'Brien S, Barrientos JC, Kay NE, et al. (2014) Ibrutinib versus ofatumumab in previously treated chronic lymphoid leukemia. *N Engl J Med* 371: 213-223.
220. Follows GA, Bloor A, Dearden C, Devereux S, Fox CP, et al. (2015) Interim statement from the BCSH CLL Guidelines Panel.
221. National Institute for Health and Care Excellence (2007) Fludarabine monotherapy for the first-line treatment of chronic lymphocytic leukaemia.
222. National Institute for Health and Care Excellence (2009) Rituximab for the first-line treatment of chronic lymphocytic leukaemia.
223. National Institute for Health and Care Excellence (2010) Rituximab for the treatment of relapsed/refractory chronic lymphocytic leukaemia.
224. National Institute for Health and Care Excellence (2011) Bendamustine for the first-line treatment of chronic lymphocytic leukaemia.
225. Clifford R, Louis T, Robbe P, Ackroyd S, Burns A, et al. (2014) SAMHD1 is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA damage. *Blood* 123: 1021-1031.
226. Clifford RM, Robbe P, Weller S, Timbs AT, Titsias M, et al. (2014) Towards Response Prediction Using Integrated Genomics in Chronic Lymphocytic Leukaemia: Results on 250 First-Line FCR Treated Patients from UK Clinical Trials. *Blood* 124: 1942-1942.
227. Byrd JC, Furman RR, Coutre SE, Burger JA, Blum KA, et al. (2015) Three-year follow-up of treatment-naïve and previously treated patients with CLL and SLL receiving single-agent ibrutinib. *Blood* 125: 2497-2506.

228. Farooqui MZ, Valdez J, Martyr S, Aue G, Saba N, et al. (2015) Ibrutinib for previously untreated and relapsed or refractory chronic lymphocytic leukaemia with TP53 aberrations: a phase 2, single-arm trial. *Lancet Oncol* 16: 169-176.
229. O'Brien S, Furman RR, Coutre SE, Sharman JP, Burger JA, et al. (2014) Ibrutinib as initial therapy for elderly patients with chronic lymphocytic leukaemia or small lymphocytic lymphoma: an open-label, multicentre, phase 1b/2 trial. *Lancet Oncology* 15: 48-58.
230. Hallek M, Fischer K, Fingerle-Rowson G, Fink AM, Busch R, et al. (2010) Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial. *Lancet* 376: 1164-1174.
231. Office for National Statistics (2013) National Life Tables, United Kingdom, 2010-2012.
232. Department of Health (2014) NHS Reference Costs 2012-2013.
233. Department of Health (2014) Drugs and pharmaceutical electronic market information (eMit).
234. Joint Formulary Committee (2014) British National Formulary (online). London: BMJ Group and Pharmaceutical Press.
235. Curtis L (2014) Unit Costs of Health & Social Care. Personal Social Services Research Unit, University of Kent.
236. Comer B (2013) Ibrutinib's Breakthrough to Market. <http://www.pharmexec.com/ibrutinibs-breakthrough-market>.
237. Roche Products Ltd (2009) Rituximab for the treatment of relapsed or refractory chronic lymphocytic leukaemia.
238. Napp Pharmaceuticals Ltd (2010) Bendamustine for the first-line treatment of chronic lymphocytic leukaemia (Binet stage B or C) in patients for whom fludarabine combination chemotherapy is not appropriate.
239. Beusterien KM, Davies J, Leach M, Meiklejohn D, Grinspan JL, et al. (2010) Population preference values for treatment outcomes in chronic lymphocytic leukaemia: a cross-sectional utility study. *Health Qual Life Outcomes* 8: 50.
240. Microsoft (2013) Microsoft Excel 2013. Redmond, Washington: Microsoft.
241. Briggs AH (2000) Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics* 17: 479-500.
242. Fenwick E, Byford S (2005) A guide to cost-effectiveness acceptability curves. *The British Journal of Psychiatry* 187: 106-108.
243. Dobes L (2008) A Century of Australian Cost-Benefit Analysis. Office of Best Practice Regulation.

244. Australian Government (2014) Independent cost-benefit analysis of broadband and review of regulation: Volume II – The costs and benefits of high-speed broadband.
245. U.S. Department of the Interior, U.S. Department of Commerce, National Marine Fisheries Service (2012) Klamath Dam Removal Overview Report for the Secretary of the Interior: An Assessment of Science and Technical Information.
246. Krol M, Brouwer W (2014) How to Estimate Productivity Costs in Economic Evaluations. *Pharmacoeconomics* 32: 335-344.
247. Clarke PM (1998) Cost-benefit analysis and mammographic screening: a travel cost approach. *J Health Econ* 17: 767-787.
248. Sach TH, Smith RD, Whynes DK (2007) A 'league table' of contingent valuation results for pharmaceutical interventions: a hard pill to swallow? *Pharmacoeconomics* 25: 107-127.
249. Louviere JJ, Flynn TN, Carson RT (2010) Discrete Choice Experiments Are Not Conjoint Analysis. *Journal of Choice Modelling* 3: 57-72.
250. Finn A, Louviere JJ (1992) Determining the Appropriate Response to Evidence of Public Concern - the Case of Food Safety. *Journal of Public Policy & Marketing* 11: 12-25.
251. Louviere J, Lings I, Islam T, Gudergan S, Flynn T (2013) An introduction to the application of (case 1) best-worst scaling in marketing research. *International Journal of Research in Marketing* 30: 292-303.
252. National Centre for Environmental Economics, Office of Policy, U.S. Environmental Protection Agency (2014) Guidelines for Preparing Economic Analyses.
253. HM Treasury (2013) The Green Book: appraisal and evaluation in central government.
254. Regier DA (2008) Discrete choice experiments informing cost benefit analysis: a Bayesian approach with an application to genetic testing: University of Aberdeen.
255. de Bekker-Grob EW, Ryan M, Gerard K (2012) Discrete choice experiments in health economics: a review of the literature. *Health Econ* 21: 145-172.
256. Payne K, Fargher EA, Roberts SA, Tricker K, Elliott RA, et al. (2011) Valuing pharmacogenetic testing services: a comparison of patients' and health care professionals' preferences. *Value Health* 14: 121-134.
257. Chan SL, Low JJW, Lim YW, Finkelstein E, Farooqui MA, et al. (2013) Willingness-to-pay and preferences for warfarin pharmacogenetic testing in Chinese warfarin patients and the Chinese general public. *Personalized Medicine* 10: 127-137.
258. Regier DA, Friedman JM, Makela N, Ryan M, Marra CA (2009) Valuing the benefit of diagnostic testing for genetic causes of idiopathic developmental disability: willingness to pay from families of affected children. *Clin Genet* 75: 514-521.

259. Cameron LD, Sherman KA, Marteau TM, Brown PM (2009) Impact of genetic risk information and type of disease on perceived risk, anticipated affect, and expected consequences of genetic tests. *Health Psychol* 28: 307-316.
260. Ryan M, Watson V, Entwistle V (2009) Rationalising the 'irrational': a think aloud study of discrete choice experiment responses. *Health Economics* 18: 321-336.
261. Schapira MM, Nattinger AB, McAuliffe TL (2006) The influence of graphic format on breast cancer risk communication. *J Health Commun* 11: 569-582.
262. Weinstein ND (1999) What Does It Mean to Understand a Risk? Evaluating Risk Comprehension. *JNCI Monographs* 1999: 15-20.
263. Harrison M, Rigby D, Vass C, Flynn T, Louviere J, et al. (2014) Risk as an attribute in discrete choice experiments: a systematic review of the literature. *Patient* 7: 151-170.
264. Donnalley L, Reed Johnson F, Gonzalez J (2012) Evolution of the Humanoid: Explaining Probabilistic Outcomes to Innumerate Conjoint Subjects.
265. Lloyd AJ (2001) The extent of patients' understanding of the risk of treatments. *Quality in Health Care* 10: i14-i18.
266. Reyna VF (2008) A theory of medical decision making and health: fuzzy trace theory. *Med Decis Making* 28: 850-865.
267. Gigerenzer G (2002) *Reckoning with Risk: Learning to Live with Uncertainty*: Penguin.
268. Gigerenzer G, Galesic M (2012) Why do single event probabilities confuse patients? *BMJ* 344.
269. Hensher D, Greene W (2010) Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification. *Empirical Economics* 39: 413-426.
270. Cameron LD, Marteau TM, Brown PM, Klein WM, Sherman KA (2012) Communication strategies for enhancing understanding of the behavioral implications of genetic and biomarker tests for disease risk: the role of coherence. *J Behav Med* 35: 286-298.
271. Sung L, Alibhai SM, Ethier MC, Teuffel O, Cheng S, et al. (2012) Discrete choice experiment produced estimates of acceptable risks of therapeutic options in cancer patients with febrile neutropenia. *J Clin Epidemiol* 65: 627-634.
272. Henneman L, Oosterwijk JC, van Asperen CJ, Menko FH, Ockhuysen-Vermey CF, et al. (2013) The effectiveness of a graphical presentation in addition to a frequency format in the context of familial breast cancer risk communication: a multicenter controlled trial. *BMC Med Inform Decis Mak* 13: 55.
273. Lancsar E, Louviere J (2006) Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? *Health Economics* 15: 797-811.

274. Spiegelhalter D, Pearson M, Short I (2011) Visualizing Uncertainty About the Future. *Science* 333: 1393-1400.
275. Bech M, Kjaer T, Lauridsen J (2011) Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment. *Health Econ* 20: 273-286.
276. ChoiceMetrics (2012) Ngene 1.1.1 User Manual & Reference Guide. Australia.
277. StataCorp (2011) Stata Statistical Software: Release 12. College Station, TX: StataCorp LP.
278. Rose J, Scarpa R, Bliemer M (2008) Incorporating model uncertainty into the generation of efficient stated choice experiments: A model averaging approach. Working paper ITLS-WP-09-08, Institute of Transport and Logistics Studies.
279. Ryan M, Gerard K, Amaya-Amaya M (2008) Using discrete choice experiments to value health and health care: Springer.
280. Swait J, Louviere J (1993) The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit-Models. *Journal of Marketing Research* 30: 305-314.
281. Kind P, Hardman G, Macran S (1999) UK Population Norms for EQ-5D. University of York, Centre for Health Economics, Discussion Paper 172.
282. Sullivan PW, Slejko JF, Sculpher MJ, Ghushchyan V (2011) Catalogue of EQ-5D scores for the United Kingdom. *Med Decis Making* 31: 800-804.
283. Catovsky D, Wade R, Else M (2014) The clinical significance of patients' sex in chronic lymphocytic leukemia. *Haematologica* 99: 1088-1094.
284. Galdas PM, Cheater F, Marshall P (2005) Men and health help-seeking behaviour: literature review. *J Adv Nurs* 49: 616-623.
285. Oberoi DV, Jiwa M, McManus A, Hodder R (2014) Colorectal cancer--applying a gender lens. *Qual Prim Care* 22: 71-79.
286. Regier DA, Ryan M, Phimister E, Marra CA (2009) Bayesian and classical estimation of mixed logit: An application to genetic testing. *J Health Econ* 28: 598-610.
287. Herbild L, Bech M, Gyrd-Hansen D (2009) Estimating the Danish Populations' Preferences for Pharmacogenetic Testing Using a Discrete Choice Experiment. The Case of Treating Depression. *Value in Health* 12: 560-567.
288. Najafzadeh M, Johnston KM, Peacock SJ, Connors JM, Marra MA, et al. (2013) Genomic testing to determine drug response: measuring preferences of the public and patients using Discrete Choice Experiment (DCE). *BMC Health Services Research* 13: 454-454.
289. Watson V, Becker F, de Bekker-Grob E (2016) Discrete Choice Experiment Response Rates: A Meta-analysis. *Health Economics: n/a-n/a*.

290. Bogelund M, Vilsboll T, Faber J, Henriksen JE, Gjesing RP, et al. (2011) Patient preferences for diabetes management among people with type 2 diabetes in Denmark - a discrete choice experiment. *Curr Med Res Opin* 27: 2175-2183.
291. Marshall D, Bridges JF, Hauber B, Cameron R, Donnalley L, et al. (2010) Conjoint Analysis Applications in Health - How are Studies being Designed and Reported?: An Update on Current Practice in the Published Literature between 2005 and 2008. *Patient* 3: 249-256.
292. Wanders JO, Veldwijk J, de Wit GA, Hart HE, van Gils PF, et al. (2014) The effect of out-of-pocket costs and financial rewards in a discrete choice experiment: an application to lifestyle programs. *BMC Public Health* 14: 870.
293. Powell G, Holmes EAF, Plumpton CO, Ring A, Baker GA, et al. (2015) Pharmacogenetic testing prior to carbamazepine treatment of epilepsy: patients' and physicians' preferences for testing and service delivery. *British Journal of Clinical Pharmacology* 80: 1149-1159.
294. Tan SS, Bouwmans CA, Rutten FF, Hakkaart-van Roijen L (2012) Update of the Dutch Manual for Costing in Economic Evaluations. *Int J Technol Assess Health Care* 28: 152-158.
295. Blankart CR, Koch T, Linder R, Verheyen F, Schreyögg J, et al. (2013) Cost of illness and economic burden of chronic lymphocytic leukemia. *Orphanet Journal of Rare Diseases* 8: 32-32.
296. Finney A, Davies S, Hayes D, Collard S (2013) Cancer's hidden price tag: Revealing the costs behind the illness. *Macmillan Cancer Support*.
297. Claxton KP, Sculpher MJ (2006) Using Value of Information Analysis to Prioritise Health Research. *Pharmacoeconomics* 24: 1055-1068.
298. Watson L, Wyld P, Catovsky D (2008) Disease burden of chronic lymphocytic leukaemia within the European Union. *Eur J Haematol* 81: 253-258.
299. Office for National Statistics (2013) Annual Mid-year Population Estimates.
300. Malhotra C, Farooqui MA, Kanesvaran R, Bilger M, Finkelstein E (2015) Comparison of preferences for end-of-life care among patients with advanced cancer and their caregivers: A discrete choice experiment. *Palliative Medicine*.
301. Lloyd A, Penson D, Dewilde S, Kleinman L (2008) Eliciting patient preferences for hormonal therapy options in the treatment of metastatic prostate cancer. *Prostate Cancer Prostatic Dis* 11: 153-159.
302. Robak T, Dmoszynska A, Solal-Celigny P, Warzocha K, Loscertales J, et al. (2010) Rituximab plus fludarabine and cyclophosphamide prolongs progression-free survival compared with fludarabine and cyclophosphamide alone in previously treated chronic lymphocytic leukemia. *J Clin Oncol* 28: 1756-1765.

303. Badoux XC, Keating MJ, Wang X, O'Brien SM, Ferrajoli A, et al. (2011) Fludarabine, cyclophosphamide, and rituximab chemoimmunotherapy is highly effective treatment for relapsed patients with CLL. *Blood* 117: 3016-3024.
304. Abrisqueta P, Pereira A, Rozman C, Aymerich M, Gine E, et al. (2009) Improving survival in patients with chronic lymphocytic leukemia (1980-2008): the Hospital Clinic of Barcelona experience. *Blood* 114: 2044-2050.
305. Holtzer-Goor KM, Schaafsma MR, Joosten P, Posthuma EF, Wittebol S, et al. (2015) Quality of life of patients with chronic lymphocytic leukaemia in the Netherlands: results of a longitudinal multicentre study. *Qual Life Res*.
306. Devlin NJ, Krabbe PF (2013) The development of new research methods for the valuation of EQ-5D-5L. *Eur J Health Econ* 14 Suppl 1: S1-3.
307. Goldstein DA, Shaib WL, Flowers CR (2015) Costs and effectiveness of genomic testing in the management of colorectal cancer. *Oncology (Williston Park)* 29: 175-183.
308. Drummond MF, Sculpher MJ, Claxton KP, Stoddart GL, Torrance GW (2015) *Methods for the Economic Evaluation of Health Care Programmes*.
309. Buchanan J, Wordsworth S (2015) Welfarism versus extra-welfarism: can the choice of economic evaluation approach impact on the adoption decisions recommended by economic evaluation studies? *Pharmacoeconomics* 33: 571-579.
310. Buchanan J, Wordsworth S, Schuh A (2013) Issues surrounding the health economic evaluation of genomic technologies. *Pharmacogenomics* 14: 1833-1847.
311. Lorgelly PK, Lorimer K, Fenwick EA, Briggs AH, Anand P (2015) Operationalising the capability approach as an outcome measure in public health: The development of the OCAP-18. *Soc Sci Med* 142: 68-81.
312. Brennan V, Dixon S (2013) Incorporating Process Utility into Quality Adjusted Life Years: A Systematic Review of Empirical Studies. *Pharmacoeconomics* 31: 677-691.
313. Schering Health Care Ltd (2006) Fludarabine phosphate for the first-line treatment of chronic lymphocytic leukaemia.
314. Walker S, Palmer S, Erhorn S, Brent S, Dyker A, et al. (2009) Fludarabine phosphate for the first-line treatment of chronic lymphocytic leukaemia: the evidence review group's report. National Institute for Health and Care Excellence.
315. Roche Products Ltd (2008) Rituximab for the first-line treatment of chronic lymphocytic leukaemia.
316. Main C, Pitt M, Moxham T, Stein K (2009) The clinical and cost-effectiveness of Rituximab for the 1st line treatment of Chronic Lymphocytic Leukaemia: an evidence review of the submission from Roche. National Institute for Health and Care Excellence.

317. West Midlands Health Technology Assessment Collaboration (2009) Rituximab for the treatment of relapsed or refractory chronic lymphocytic leukaemia: evidence review group report. National Institute for Health and Care Excellence.
318. GlaxoSmithKline UK (2010) Ofatumumab (Arzerra) for the treatment of chronic lymphocytic leukaemia in patients who are refractory to fludarabine and alemtuzumab.
319. Hoyle M, Crathorne L, Moxham T, Garside R, Hyde C (2010) Ofatumumab (Arzerra) for the treatment of chronic lymphocytic leukaemia in patients who are refractory to fludarabine and alemtuzumab: a critique of the submission from GSK National Institute for Health and Care Excellence.
320. Hoyle M, Crathorne L, Jones-Hughes T, Stein K (2010) Bendamustine for the first-line treatment of chronic lymphocytic leukaemia (Binet stage B or C) in patients for whom fludarabine combination chemotherapy is not appropriate: a critique of the submission from Napp. National Institute for Health and Care Excellence.
321. Shanafelt TD, Witzig TE, Fink SR, Jenkins RB, Paternoster SF, et al. (2006) Prospective Evaluation of Clonal Evolution During Long-Term Follow-Up of Patients With Untreated Early-Stage Chronic Lymphocytic Leukemia. *Journal of Clinical Oncology* 24: 4634-4641.
322. Gunnarsson R, Mansouri L, Isaksson A, Göransson H, Cahill N, et al. (2011) Array-based genomic screening at diagnosis and during follow-up in chronic lymphocytic leukemia. *Blood* 96: 1161-1169.
323. Cuneo A, Marchetti M, Barosi G, Billio A, Brugiattelli M, et al. (2014) Appropriate use of bendamustine in first-line therapy of chronic lymphocytic leukemia. Recommendations from SIE, SIES, GITMO Group. *Leukemia Research*.
324. Eichhorst B, Fink A-M, Busch R, Lange E, Köppler H, et al. (2013) Chemoimmunotherapy With Fludarabine (F), Cyclophosphamide (C), and Rituximab (R) (FCR) Versus Bendamustine and Rituximab (BR) In Previously Untreated and Physically Fit Patients (pts) With Advanced Chronic Lymphocytic Leukemia (CLL): Results Of a Planned.... *Blood* 122: 526-526.
325. Fischer K, Cramer P, Busch R, Bottcher S, Bahlo J, et al. (2012) Bendamustine in combination with rituximab for previously untreated patients with chronic lymphocytic leukemia: a multicenter phase II trial of the German Chronic Lymphocytic Leukemia Study Group. *J Clin Oncol* 30: 3209-3216.
326. Eketorp Sylvan S, Hansson L, Karlsson C, Norin S, Lundin J, et al. (2014) Outcomes of patients with fludarabine-refractory chronic lymphocytic leukemia: a population-based study from a well-defined geographic region. *Leuk Lymphoma* 55: 1774-1780.
327. Tam CS, O'Brien S, Lerner S, Khouri I, Ferrajoli A, et al. (2007) The natural history of fludarabine-refractory chronic lymphocytic leukemia patients who fail alemtuzumab or have bulky lymphadenopathy. *Leuk Lymphoma* 48: 1931-1939.

328. Fischer K, Cramer P, Busch R, Stilgenbauer S, Bahlo J, et al. (2011) Bendamustine Combined With Rituximab in Patients With Relapsed and/or Refractory Chronic Lymphocytic Leukemia: A Multicenter Phase II Trial of the German Chronic Lymphocytic Leukemia Study Group. *Journal of Clinical Oncology* 29: 3559-3566.
329. Wierda WG, Kipps TJ, Mayer J, Stilgenbauer S, Williams CD, et al. (2010) Ofatumumab as single-agent CD20 immunotherapy in fludarabine-refractory chronic lymphocytic leukemia. *J Clin Oncol* 28: 1749-1755.
330. Dreger P, Döhner H, Ritgen M, Böttcher S, Busch R, et al. (2010) Allogeneic stem cell transplantation provides durable disease control in poor-risk chronic lymphocytic leukemia: long-term clinical and MRD results of the German CLL Study Group CLL3X trial. *Blood* 116: 2438-2447.
331. Wierda W, O'Brien S, Wen S, Faderl S, Garcia-Manero G, et al. (2005) Chemoimmunotherapy with fludarabine, cyclophosphamide, and rituximab for relapsed and refractory chronic lymphocytic leukemia. *J Clin Oncol* 23: 4070-4078.
332. Marks DI, Lush R, Cavenagh J, Milligan DW, Schey S, et al. (2002) The toxicity and efficacy of donor lymphocyte infusions given after reduced-intensity conditioning allogeneic stem cell transplantation. *Blood* 100: 3108-3114.
333. Punekar YS, Shukla A, Mullerova H (2014) COPD management costs according to the frequency of COPD exacerbations in UK primary care. *Int J Chron Obstruct Pulmon Dis* 9: 65-73.
334. Department of Health (2013) NHS Blood and Transplant Price List 2012/13.
335. Hoyle M, Pavey T, Ciani O, Crathorne L, Jones-Hughes T, et al. (2011) Dasatinib, nilotinib and standard dose imatinib for the first-line treatment of chronic myeloid leukaemia: systematic reviews and economic analyses. University of Exeter (Report for NICE).
336. Health and Social Care Information Centre (2013) Health Survey for England - 2012.
337. Hancock S, Wake B, Hyde C (2003) Fludarabine as first line therapy for chronic lymphocytic leukaemia. West Midlands Health Technology Assessment Collaboration.
338. Kharfan-Dabaja MA, Pidala J, Kumar A, Terasawa T, Djulbegovic B (2012) Comparing efficacy of reduced-toxicity allogeneic hematopoietic cell transplantation with conventional chemo-(immuno) therapy in patients with relapsed or refractory CLL: a Markov decision analysis. *Bone Marrow Transplant* 47: 1164-1170.
339. Ferguson J, Tolley K, Gilmour L, Priaux J (2008) Health state preference study mapping the change over the course of the disease process in chronic lymphocytic leukaemia (CLL). *Value in Health* 11: A485.

340. Tolley K, Goad C, Yi Y, Maroudas P, Haiderali A, et al. (2013) Utility elicitation study in the UK general public for late-stage chronic lymphocytic leukaemia. *Eur J Health Econ* 14: 749-759.
341. Armstrong K, Putt M, Hughes halbert C, Grande D, Shwartz JS, et al. (2011) PCN158 The Influence of Health Care Policies and Health Care System Distrust on Willingness to Undergo Genetic Testing. *Value in Health* 14: A463.
342. Bishop AJ, Marteau TM, Armstrong D, Chitty LS, Longworth L, et al. (2004) Women and health care professionals' preferences for Down's Syndrome screening tests: a conjoint analysis study. *BJOG* 111: 775-779.
343. Cairns J, Shackley P, Hundley V (1996) Decision making with respect to diagnostic testing: a method of valuing the benefits of antenatal screening. *Med Decis Making* 16: 161-168.
344. Griffith G, Edwards R, Williams J, Gray J, Morrison V, et al. (2009) Patient preferences and National Health Service costs: a cost-consequences analysis of cancer genetic services. *Familial Cancer* 8: 265-275.
345. Hall J, Fiebig DG, King MT, Hossain I, Louviere JJ (2006) What influences participation in genetic carrier testing? Results from a discrete choice experiment. *J Health Econ* 25: 520-537.
346. Herbild L (2006) Pharmacogenomic screening for CYP-polymorphisms in the treatment of depression - economics and patient preferences. *Economics of Genetic Technologies Seminar Series*. Manchester.
347. Herbild L, Gyrd-Hansen D, Bech M (2008) Patient preferences for pharmacogenetic screening in depression. *Int J Technol Assess Health Care* 24: 96-103.
348. Huang MY, Perri M (2011) PND50 Public Preferences for the Predictive Genetic Test for Alzheimer'S Disease in the United States. *Value in Health* 14: A326.
349. Lee JT, Bridges JF, Shockney L (2008) Can pharmacoeconomics and outcomes research contribute to the empowerment of women affected by breast cancer? *Expert Rev Pharmacoecon Outcomes Res* 8: 73-79.
350. Lewis SM, Cullinane FM, Carlin JB, Halliday JL (2006) Women's and health professionals' preferences for prenatal testing for Down syndrome in Australia. *Australian and New Zealand Journal of Obstetrics and Gynaecology* 46: 205-211.
351. Marshall DA, Johnson FR, Phillips KA, Marshall JK, Thabane L, et al. (2007) Measuring Patient Preferences for Colorectal Cancer Screening Using a Choice-Format Survey. *Value in Health* 10: 415-430.
352. Mohamed AF, Johnson FR, Knight SJ (2011) Preferences for genetic testing to assess colon-cancer risks. *International Meeting of the International Society for Pharmacoeconomics and Outcomes Research*. Baltimore, MD United States.

353. Peacock S, Apicella C, Andrews L, Tucker K, Bankier A, et al. (2006) A discrete choice experiment of preferences for genetic counselling among Jewish women seeking cancer genetics services. *Br J Cancer* 95: 1448-1453.
354. Reed Johnson F, Van Houtven G, Özdemir S, Hass S, White J, et al. (2009) Multiple sclerosis patients—benefit-risk preferences: Serious adverse event risks versus treatment efficacy. *Journal of Neurology* 256: 554-562.
355. Ryan M, Diack J, Watson V, Smith N (2005) Rapid prenatal diagnostic testing for Down syndrome only or longer wait for full karyotype: the views of pregnant women. *Prenatal Diagnosis* 25: 1206-1211.
356. Sebban C, Browman G, Gafni A, Norman G, Levine M, et al. (1995) Design and validation of a bedside decision instrument to elicit a patient's preference concerning allogeneic bone marrow transplantation in chronic myeloid leukemia. *Am J Hematol* 48: 221-227.
357. Severin F, Schmidtke J, Muhlbacher A, Rogowski WH (2013) Eliciting preferences for priority setting in genetic testing: a pilot study comparing best-worst scaling and discrete-choice experiments. *Eur J Hum Genet* 21: 1202-1208.
358. Guest G, Bunce A, Johnson L (2006) How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods* 18: 59-82.
359. Krol M, Brouwer WB, Severens JL, Kaper J, Evers SM (2012) Productivity cost calculations in health economic evaluations: correcting for compensation mechanisms and multiplier effects. *Soc Sci Med* 75: 1981-1988.
360. Department for Work & Pensions (2013) Older Workers Statistical Information Booklet 2013: Official Statistics.
361. Office for National Statistics (2014) Patterns of Pay: Estimates from the Annual Survey of Hours and Earnings, UK, 1997 to 2013.
362. Luengo-Fernandez R, Leal J, Gray A, Sullivan R (2013) Economic burden of cancer across the European Union: a population-based cost analysis. *Lancet Oncol* 14: 1165-1174.
363. Luengo-Fernández R, Leal J, Gray A, Petersen S, Rayner M (2006) Cost of cardiovascular diseases in the United Kingdom. *Heart* 92: 1384-1389.
364. Krol M (2012) Productivity costs in economic evaluations: Erasmus University Rotterdam,.
365. www.gov.uk (2015) National Minimum Wage rates.
366. Schultz A, Chen C-Y, Edington D (2009) The Cost and Impact of Health Conditions on Presenteeism to Employers. *Pharmacoeconomics* 27: 365-378.

367. Goetzel RZ, Long SR, Ozminkowski RJ, Hawkins K, Wang S, et al. (2004) Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting U.S. employers. *J Occup Environ Med* 46: 398-412.
368. Krol M, Brouwer W, Sendi P (2006) Productivity costs in health-state valuations : does explicit instruction matter? *Pharmacoeconomics* 24: 401-414.
369. Krol M, Sendi P, Brouwer W (2009) Breaking the silence: exploring the potential effects of explicit instructions on incorporating income and leisure in TTO exercises. *Value Health* 12: 172-180.
370. Tilling C, Krol M, Tsuchiya A, Brazier J, van Exel J, et al. (2012) Does the EQ-5D Reflect Lost Earnings? *Pharmacoeconomics* 30: 47-61.
371. Burns R, Luengo-Fernandez R, Leal J (2015) Economic Burden of Disorders of the Blood in Europe. European Hematology Association 20th Congress. Vienna, Austria.
372. Van Houtven CH, Ramsey SD, Hornbrook MC, Atienza AA, van Ryn M (2010) Economic Burden for Informal Caregivers of Lung and Colorectal Cancer Patients. *The Oncologist* 15: 883-893.

Appendix 1: Search strategies used to identify references which consider the impact of choice of economic evaluation approach on adoption decisions

Embase (1988 to present)

cost effectiveness. ti,ab.
cost utility. ti,ab.
cost minimi#ation. ti,ab.
cost consequence.ti,ab.
Or/1-4
cost benefit. ti,ab.
5 and 6
limit 7 to english language
editorial.pt
letter.pt
note.pt
(rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep or primate or primates).ti,ab.
or/9-12
8 not 13
limit 14 to human
limit 15 to yr="1990 -Current"

Medline In-Process & Other Non-Indexed Citations and Medline 1946 to present + Medline Daily Update

cost effectiveness. ti,ab.
cost utility. ti,ab.
cost minimi#ation. ti,ab.
cost consequence.ti,ab.
Or/1-4
cost benefit. ti,ab.
5 and 6
limit 7 to english language
editorial.pt
letter.pt
(rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep or primate or primates).ti,ab.
or/9-11
8 not 12
limit 13 to human
limit 14 to yr="1990 -Current"

Econlit

(ti,ab("cost effectiveness") OR ti,ab("cost utility") OR ti,ab("cost consequence") OR ti,ab("cost minimi?ation")) AND ti,ab("cost benefit")

Limited to English language, excluding dissertations, from 1990 until 22/08/13

The Cochrane Library

Searched for the following in title, abstract and keywords (combined with AND):

cost utility or cost effectiveness or cost minimisation or cost consequence [title, abstract, keywords]

cost benefit [record title]

Centre for Reviews and Dissemination

Searched for the following across all 3 databases (combined with AND):

cost utility or cost effectiveness or cost minimisation or cost consequence [any field]

cost benefit [title]

Date limited to 1990/01/01-2013/08/22

Note

As this literature review was designed to identify references within some broad categories, several steps were taken to ensure that the search was as specific as possible, whilst still minimising false negatives. Within Embase and Medline certain Medical Subject Headings were not expanded (e.g. “health economics” or “economic evaluation”) as these categories were too broad and there was no sensible way to reduce the number of hits to a manageable level. It was judged that using different combinations of terms for the different forms of economic evaluation would be a more effective way to approach the review. In the Cochrane Library search the phrase ‘cost-benefit’ was searched for in the record title only. This is a key search term but searching for it across title, abstract, and keywords (the only other option) returns an unmanageable number of results. Similarly, in the Centre for Reviews and Dissemination database the phrase ‘cost-benefit’ was searched for in the title only, rather than ‘any field’.

Appendix 2: Consolidated Health Economic Evaluation Reporting Standards checklist and quality assessment of papers included in the literature review

	Item	Recommendation
1	Title	Identify the study as an economic evaluation or use more specific terms such as “cost-effectiveness analysis”, and describe the interventions compared.
2	Abstract	Provide a structured summary of objectives, perspective, setting, methods (including study design and inputs), results (including base case and uncertainty analyses), and conclusions.
3	Background and objectives	Provide an explicit statement of the broader context for the study. Present the study question and its relevance for health policy or practice decisions.
4	Target population and subgroups	Describe characteristics of the base case population and subgroups analysed, including why they were chosen.
5	Setting and location	State relevant aspects of the system(s) in which the decision(s) need(s) to be made.
6	Study perspective	Describe the perspective of the study and relate this to the costs being evaluated.
7	Comparators	Describe the interventions or strategies being compared and state why they were chosen.
8	Time horizon	State the time horizon(s) over which costs and consequences are being evaluated and say why appropriate.
9	Discount rate	Report the choice of discount rate(s) used for costs and outcomes and say why appropriate.
10	Choice of health outcomes	Describe what outcomes were used as the measure(s) of benefit in the evaluation and their relevance for the type of analysis performed.
11	Measurement of effectiveness	<i>Single study-based estimates:</i> Describe fully the design features of the single effectiveness study and why the single study was a sufficient source of clinical effectiveness data. <i>Synthesis-based estimates:</i> Describe fully the methods used for identification of included studies and synthesis of clinical effectiveness data.
12	Measurement and valuation of preference based outcomes	If applicable, describe the population and methods used to elicit preferences for outcomes.
13	Estimating resources and costs	<i>Single study-based economic evaluation:</i> Describe approaches used to estimate resource use associated with the alternative interventions. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs. <i>Model-based economic evaluation:</i> Describe approaches and data sources used to estimate resource use associated with model health states. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.
14	Currency, price date, and conversion	Report the dates of the estimated resource quantities and unit costs. Describe methods for adjusting estimated unit costs to the year of reported costs if necessary. Describe methods for converting costs into a common currency base and the exchange rate.

	Item	Recommendation
15	Choice of model	Describe and give reasons for the specific type of decision-analytical model used. Providing a figure to show model structure is strongly recommended.
16	Assumptions	Describe all structural or other assumptions underpinning the decision-analytical model.
17	Analytical methods	Describe all analytical methods supporting the evaluation. This could include methods for dealing with skewed, missing, or censored data; extrapolation methods; methods for pooling data; approaches to validate or make adjustments (such as half cycle corrections) to a model; and methods for handling population heterogeneity and uncertainty.
18	Study parameters	Report the values, ranges, references, and, if used, probability distributions for all parameters. Report reasons or sources for distributions used to represent uncertainty where appropriate. Providing a table to show the input values is strongly recommended.
19	Incremental costs and outcomes	For each intervention, report mean values for the main categories of estimated costs and outcomes of interest, as well as mean differences between the comparator groups. If applicable, report incremental cost-effectiveness ratios.
20	Characterising uncertainty	<i>Single study-based economic evaluation:</i> Describe the effects of sampling uncertainty for the estimated incremental cost and incremental effectiveness parameters, together with the impact of methodological assumptions (such as discount rate, study perspective). <i>Model-based economic evaluation:</i> Describe the effects on the results of uncertainty for all input parameters, and uncertainty related to the structure of the model and assumptions.
21	Characterising heterogeneity	If applicable, report differences in costs, outcomes, or cost-effectiveness that can be explained by variations between subgroups of patients with different baseline characteristics or other observed variability in effects that are not reducible by more information.
22	Study findings, limitations, generalisability, and current knowledge	Summarise key study findings and describe how they support the conclusions reached. Discuss limitations and the generalisability of the findings and how the findings fit with current knowledge.
23	Source of funding	Describe how the study was funded and the role of the funder in the identification, design, conduct, and reporting of the analysis. Describe other non-monetary sources of support.
24	Conflicts of interest	Describe any potential for conflict of interest of study contributors in accordance with journal policy. In the absence of a journal policy, we recommend authors comply with International Committee of Medical Journal Editors recommendations.

Study 1: Andresen (2010)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of uncertainty
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Introduction	1	-
5	Setting and location	-	0	Aspects of system not stated
6	Study perspective	Abstract	1	-
7	Comparators	-	0	Unclear
8	Time horizon	Methods	0.5	Not justified
9	Discount rate	-	0	-
10	Choice of health outcomes	Methods	0.5	Relevance not discussed
11	Measurement of effectiveness	Methods	0.5	Methods for identification not stated. Not synthesised
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	0.5	Only partially described
14	Currency, price date, and conversion	Methods	0	-
15	Choice of model	Methods	0.5	Lacking detailed description and figure
16	Assumptions	Methods	1	-
17	Analytical methods	-	0	-
18	Study parameters	Methods	0.5	Some detail in text but not tabulated
19	Incremental costs and outcomes	Results	0.5	Means provided but incremental cost-effectiveness ratios unclear
20	Characterising uncertainty	Results	0.5	Not systematic. Structure not considered
21	Characterising heterogeneity	-	0	-
22	Study findings, limitations, generalisability, and current knowledge	Interpretation	1	-
23	Source of funding	-	0	-
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			10.5 (46%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 2: Beale (2012)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	0.5	Economic evaluation not specified
2	Abstract	Abstract	1	-
3	Background and objectives	Background	1	-
4	Target population and subgroups	-	0	Only partial information provided
5	Setting and location	Title/background	1	-
6	Study perspective	Methods	1	-
7	Comparators	Methods	1	-
8	Time horizon	Methods	1	-
9	Discount rate	Methods	0.5	No justification provided
10	Choice of health outcomes	Introduction/methods	1	-
11	Measurement of effectiveness	Methods	0.5	Methods for identification not stated. Synthesis unclear
12	Measurement and valuation of preference based outcomes	Methods	0.5	Only partially completed
13	Estimating resources and costs	Methods	0.5	Data not presented in full
14	Currency, price date, and conversion	Methods	1	-
15	Choice of model	-	0	-
16	Assumptions	Methods	1	-
17	Analytical methods	-	0	-
18	Study parameters	-	0	-
19	Incremental costs and outcomes	Results	0.5	Only partially completed
20	Characterising uncertainty	Results	0.5	Only partially completed
21	Characterising heterogeneity	-	0	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	End of paper	0.5	No information on role of funder
24	Conflicts of interest	End of paper	1	-
TOTAL SCORE (% of max possible)			15 (63%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 3: Brisson (2006)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	0.5	Economic evaluation not specified
2	Abstract	Abstract	1	-
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Methods	1	-
5	Setting and location	Methods	0.5	Regulatory setting not described
6	Study perspective	Methods	1	-
7	Comparators	Methods	1	-
8	Time horizon	Methods	1	-
9	Discount rate	Methods	1	-
10	Choice of health outcomes	Introduction/methods	1	-
11	Measurement of effectiveness	Methods	0.5	-
12	Measurement and valuation of preference based outcomes	Methods	0.5	-
13	Estimating resources and costs	Methods	0.5	-
14	Currency, price date, and conversion	Methods	0.5	-
15	Choice of model	Methods	1	-
16	Assumptions	Methods	1	-
17	Analytical methods	Methods	1	-
18	Study parameters	Appendix	1	-
19	Incremental costs and outcomes	Earlier study	0.5	-
20	Characterising uncertainty	Results	1	-
21	Characterising heterogeneity	Earlier study	1	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	Start of paper	0.5	No information on role of funder
24	Conflicts of interest	Start of paper	1	-
TOTAL SCORE (% of max possible)			20 (83%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 4: Burr (2012)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	1	-
3	Background and objectives	Earlier chapters in report	1	-
4	Target population and subgroups	Methods	1	-
5	Setting and location	Methods	1	-
6	Study perspective	Methods	1	-
7	Comparators	Methods	1	-
8	Time horizon	Methods	1	-
9	Discount rate	Methods	1	-
10	Choice of health outcomes	Methods	1	-
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	Methods	1	-
13	Estimating resources and costs	Methods	0.5	Limited information provided on resource use
14	Currency, price date, and conversion	Methods	0.5	No information provided on how unit costs were adjusted to the year of reported costs
15	Choice of model	Methods	1	-
16	Assumptions	Methods	1	-
17	Analytical methods	Methods	1	-
18	Study parameters	Methods	1	-
19	Incremental costs and outcomes	Methods	1	-
20	Characterising uncertainty	Results	1	-
21	Characterising heterogeneity	Results	1	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	Start of report	1	-
24	Conflicts of interest	Start of report	1	-
TOTAL SCORE (% of max possible)			23 (96%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 5: Caughey (2002)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of perspective, setting or uncertainty
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	-	0	-
5	Setting and location	Results	1	-
6	Study perspective	-	0	-
7	Comparators	Introduction/methods	1	-
8	Time horizon	-	0	-
9	Discount rate	Methods	0.5	No justification provided
10	Choice of health outcomes	Methods	0.5	Relevance not discussed
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	-	0	-
13	Estimating resources and costs	Methods	0.5	Lacking information on resource use
14	Currency, price date, and conversion	Methods	1	-
15	Choice of model	Methods	0.5	No justification provided
16	Assumptions	-	0	-
17	Analytical methods	-	0	-
18	Study parameters	Methods	1	-
19	Incremental costs and outcomes	Results	1	-
20	Characterising uncertainty	Results	0.5	Only partially completed
21	Characterising heterogeneity	-	0	-
22	Study findings, limitations, generalisability, and current knowledge	Comment	1	-
23	Source of funding	Start of paper	0.5	No information on role of funder
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			12.5 (52%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 6: Chen (1998)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of perspective, setting or methods
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Introduction	1	-
5	Setting and location	-	0	-
6	Study perspective	-	0	-
7	Comparators	Introduction	1	-
8	Time horizon	-	0	-
9	Discount rate	-	0	-
10	Choice of health outcomes	Methods	0.5	Only partially described
11	Measurement of effectiveness	-	0	-
12	Measurement and valuation of preference based outcomes	-	0	-
13	Estimating resources and costs	Methods	0.5	Only partially described
14	Currency, price date, and conversion	Methods	0.5	Only partially described
15	Choice of model	Methods	0.5	Only partially described. No justification provided
16	Assumptions	-	0	-
17	Analytical methods	-	0	-
18	Study parameters	Methods/results	0.5	Limited data on utilities and effectiveness
19	Incremental costs and outcomes	Results	0.5	No disaggregation.
20	Characterising uncertainty	Results	0.5	Unclear whether all parameters were varied
21	Characterising heterogeneity	Results	1	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	0.5	Lack of discussion of limitations, generalisability and current knowledge
23	Source of funding	-	0	-
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			9.5 (40%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 7: Driessen (2012)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of setting or uncertainty
3	Background and objectives	Introduction	0.5	Limited context or consideration of policy relevance
4	Target population and subgroups	Methods	0.5	No justification provided
5	Setting and location	Methods	0.5	Relevant aspects of system not described
6	Study perspective	Methods	1	-
7	Comparators	Introduction	1	-
8	Time horizon	-	0	-
9	Discount rate	-	0	-
10	Choice of health outcomes	Methods	1	-
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	1	-
14	Currency, price date, and conversion	Methods	0.5	Unclear
15	Choice of model	-	N/A	-
16	Assumptions	-	N/A	-
17	Analytical methods	Methods	1	-
18	Study parameters	-	0	-
19	Incremental costs and outcomes	Results	0.5	Only partially reported. No disaggregation
20	Characterising uncertainty	-	0	-
21	Characterising heterogeneity	-	0	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	Earlier study	0.5	-
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			11.5 (55%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 8: Evans (1997)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of setting
3	Background and objectives	Introduction	0.5	Policy relevance not considered
4	Target population and subgroups	Methods	1	-
5	Setting and location	-	0	-
6	Study perspective	Methods	1	-
7	Comparators	Introduction	1	-
8	Time horizon	Methods	1	-
9	Discount rate	Methods	1	-
10	Choice of health outcomes	Methods	1	-
11	Measurement of effectiveness	Methods	0.5	Only partially described
12	Measurement and valuation of preference based outcomes	Methods	1	-
13	Estimating resources and costs	Methods	0.5	Lack of information on resource use
14	Currency, price date, and conversion	Methods	0.5	Dates not reported
15	Choice of model	Methods	0.5	No justification provided
16	Assumptions	Methods	1	-
17	Analytical methods	Methods	1	-
18	Study parameters	-	0	-
19	Incremental costs and outcomes	Results	1	-
20	Characterising uncertainty	Results	1	-
21	Characterising heterogeneity	Results	0.5	Limited consideration
22	Study findings, limitations, generalisability, and current knowledge	Conclusions	0.5	No consideration of how findings fit with current knowledge
23	Source of funding	End of paper	0.5	No information on role of funder
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			16.5 (69%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 9: Ginsberg (2001)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of methodology or uncertainty
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Introduction	1	-
5	Setting and location	Introduction	1	-
6	Study perspective	Methods	1	-
7	Comparators	Methods	0.5	Unclear
8	Time horizon	Methods	1	-
9	Discount rate	Methods	0.5	No justification provided
10	Choice of health outcomes	Methods	0.5	Relevance not considered
11	Measurement of effectiveness	Methods	0.5	Lack of information on synthesis
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	1	-
14	Currency, price date, and conversion	Methods	1	-
15	Choice of model	Methods	0.5	Only partially described
16	Assumptions	Methods	0.5	Only partially described
17	Analytical methods	Methods	0.5	Only partially described
18	Study parameters	Methods	0.5	No ranges or distributions provided
19	Incremental costs and outcomes	Results	0.5	Incremental cost-effectiveness ratios not directly calculated
20	Characterising uncertainty	Results	1	-
21	Characterising heterogeneity	-	N/A	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	End of paper	0.5	No information on role of funder
24	Conflicts of interest	End of paper	1	-
TOTAL SCORE (% of max possible)			16.5 (75%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 10: Goodchild (2011)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of perspective, methodology or uncertainty
3	Background and objectives	Introduction	0.5	Lacking detail
4	Target population and subgroups	Introduction	1	-
5	Setting and location	Introduction/methods	0.5	Lack of information on setting
6	Study perspective	-	0	-
7	Comparators	-	0	-
8	Time horizon	Methods	0.5	Not sufficiently specified
9	Discount rate	Methods	0.5	Only partial information provided
10	Choice of health outcomes	Methods	1	-
11	Measurement of effectiveness	Methods	0.5	Only partially described
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	0.5	Limited information provided
14	Currency, price date, and conversion	-	0	-
15	Choice of model	-	N/A	-
16	Assumptions	-	N/A	-
17	Analytical methods	Methods	0.5	Only partial information provided
18	Study parameters	Methods	0.5	Only partial information provided
19	Incremental costs and outcomes	Results	0.5	Insufficiently disaggregated
20	Characterising uncertainty	Results	0.5	Unclear whether all parameters were varied
21	Characterising heterogeneity	-	N/A	-
22	Study findings, limitations, generalisability, and current knowledge	-	0	-
23	Source of funding	End of paper	0.5	No information on role of funder
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			9 (45%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 11: Grosse (2005)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	0.5	Economic evaluation not specified
2	Abstract	Abstract	0	-
3	Background and objectives	Page 1	1	-
4	Target population and subgroups	Page 1	1	-
5	Setting and location	Page 1	1	-
6	Study perspective	-	0	-
7	Comparators	-	0	-
8	Time horizon	-	0	-
9	Discount rate	Page 4	1	-
10	Choice of health outcomes	Throughout	1	-
11	Measurement of effectiveness	Page 3	1	-
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Page 4	0.5	All relevant data not presented
14	Currency, price date, and conversion	Page 4	0.5	Only partial information provided
15	Choice of model	-	N/A	-
16	Assumptions	-	N/A	-
17	Analytical methods	-	0	-
18	Study parameters	-	0	-
19	Incremental costs and outcomes	-	0	-
20	Characterising uncertainty	Page 4	0.5	Only partially considered
21	Characterising heterogeneity	-	N/A	-
22	Study findings, limitations, generalisability, and current knowledge	Page 3	0.5	Lacking discussion of limitations and generalisability
23	Source of funding	-	0	-
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			8.5 (43%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 12: Holtgrave (1993)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	1	-
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Introduction	1	-
5	Setting and location	Introduction	1	-
6	Study perspective	Methods	0.5	Not related to costs
7	Comparators	-	0	-
8	Time horizon	Methods	0.5	Unclearly specified
9	Discount rate	Methods	1	-
10	Choice of health outcomes	Methods	1	-
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	1	-
14	Currency, price date, and conversion	Methods	0.5	Not described in full
15	Choice of model	-	N/A	-
16	Assumptions	-	N/A	-
17	Analytical methods	Methods	0.5	Only partially described
18	Study parameters	Methods	0.5	No ranges provided
19	Incremental costs and outcomes	-	0	-
20	Characterising uncertainty	Methods/results	1	-
21	Characterising heterogeneity	Results	0.5	Limited consideration
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	-	0	-
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			14 (67%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 13: Johannesson (1992)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	0.5	Economic evaluation not specified
2	Abstract	Abstract	0.5	No mention of perspective or uncertainty
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Methods	1	-
5	Setting and location	Methods	0.5	No mention of system
6	Study perspective	-	0	-
7	Comparators	Introduction/methods	1	-
8	Time horizon	Methods	1	-
9	Discount rate	Methods	0.5	No justification provided
10	Choice of health outcomes	Methods	1	-
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	1	-
14	Currency, price date, and conversion	Methods	0.5	Only partially described
15	Choice of model	Methods	0.5	Only partially described
16	Assumptions	Methods	0.5	Only partially described
17	Analytical methods	Methods	0.5	Only partially described
18	Study parameters	Methods	0.5	Only limited information provided
19	Incremental costs and outcomes	Results	1	-
20	Characterising uncertainty	Results	0.5	Only partially characterised
21	Characterising heterogeneity	-	0	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	Start of paper	0.5	No information on role of funder
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			14.5 (63%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 14: Liu (2012)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	1	-
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Methods	1	-
5	Setting and location	Methods	1	-
6	Study perspective	-	0	-
7	Comparators	Methods	1	-
8	Time horizon	Methods	0.5	No justification provided
9	Discount rate	Methods	0.5	No justification provided
10	Choice of health outcomes	Methods	1	-
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	Methods	0.5	Methodology not described
13	Estimating resources and costs	Methods	0.5	Lack of resource use data
14	Currency, price date, and conversion	Methods	0.5	Only partially described
15	Choice of model	Methods	0.5	Only partially described, and no justification provided
16	Assumptions	Methods	1	-
17	Analytical methods	Methods	0.5	Limited description provided
18	Study parameters	Methods	1	-
19	Incremental costs and outcomes	Results	1	-
20	Characterising uncertainty	Results/appendix	1	-
21	Characterising heterogeneity	-	0	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	Start of paper	0.5	No information on role of funder
24	Conflicts of interest	End of paper	1	-
TOTAL SCORE (% of max possible)			18 (75%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 15: Loisel (2002)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of uncertainty
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Methods	0.5	No information on why chosen
5	Setting and location	-	0	-
6	Study perspective	Methods	1	-
7	Comparators	Methods	0.5	No information on why chosen
8	Time horizon	Methods	0.5	No justification provided
9	Discount rate	-	0	-
10	Choice of health outcomes	Methods	0.5	Lack of explanation
11	Measurement of effectiveness	Introduction/methods	1	-
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	1	-
14	Currency, price date, and conversion	Methods	1	-
15	Choice of model	-	N/A	-
16	Assumptions	-	N/A	-
17	Analytical methods	Methods	1	-
18	Study parameters	Methods	0.5	Only partially described
19	Incremental costs and outcomes	Results	1	-
20	Characterising uncertainty	Results	0.5	Only partially characterised
21	Characterising heterogeneity	-	0	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	0.5	Generalisability and current knowledge not considered
23	Source of funding	-	0	-
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			12 (57%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 16: McCormack (2005)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	1	-
3	Background and objectives	Earlier chapters in report	1	-
4	Target population and subgroups	Methods	0.5	Limited information provided
5	Setting and location	Earlier chapters in report	1	-
6	Study perspective	Methods	1	-
7	Comparators	Methods	1	-
8	Time horizon	Methods	1	-
9	Discount rate	Methods	0.5	Not justified
10	Choice of health outcomes	Methods	0.5	Relevance not described
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	Methods	0.5	Limited information on population and methods
13	Estimating resources and costs	Methods	1	-
14	Currency, price date, and conversion	Methods	0.5	No information provided on how unit costs were adjusted to the year of reported costs
15	Choice of model	Methods	0.5	Not justified
16	Assumptions	Methods	1	-
17	Analytical methods	Methods	1	-
18	Study parameters	Methods	1	-
19	Incremental costs and outcomes	Results	1	-
20	Characterising uncertainty	Results	1	-
21	Characterising heterogeneity	Results	1	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	Start of report	1	-
24	Conflicts of interest	Start of report	1	-
TOTAL SCORE (% of max possible)			21 (88%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 17: McKay (1991)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	Only partial abstract provided
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Methods	1	-
5	Setting and location	Introduction/methods	1	-
6	Study perspective	-	0	-
7	Comparators	-	0	-
8	Time horizon	-	0	-
9	Discount rate	-	0	-
10	Choice of health outcomes	Methods	1	-
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	0.5	Unclear
14	Currency, price date, and conversion	-	0	-
15	Choice of model	-	0	-
16	Assumptions	Methods	0.5	Only partially described
17	Analytical methods	-	0	-
18	Study parameters	-	0	-
19	Incremental costs and outcomes	Results	1	-
20	Characterising uncertainty	Results	0.5	Not clear that all parameters were considered
21	Characterising heterogeneity	-	N/A	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	-	0	-
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			10 (46%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 18: Moradi-Lakeh (2012)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	1	-
3	Background and objectives	Introduction	0.5	Relevance of study question not considered
4	Target population and subgroups	Introduction	0.5	No justification provided
5	Setting and location	-	0	-
6	Study perspective	Methods	1	-
7	Comparators	Introduction	1	-
8	Time horizon	-	0	-
9	Discount rate	Methods	0.5	No justification provided
10	Choice of health outcomes	Methods	0.5	Relevance not considered
11	Measurement of effectiveness	Methods	0.5	Only limited information provided
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	0.5	Only partial descriptions provided
14	Currency, price date, and conversion	-	0	-
15	Choice of model	Methods	0.5	No justification provided
16	Assumptions	-	0	-
17	Analytical methods	-	0	-
18	Study parameters	Methods	0.5	Data missing for some parameters
19	Incremental costs and outcomes	Results	0.5	Only partially met
20	Characterising uncertainty	Results	0.5	Sensitivity analysis is not comprehensive
21	Characterising heterogeneity	-	0	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	0.5	Generalisability and current knowledge not considered
23	Source of funding	End of paper	0.5	No information on role of funder
24	Conflicts of interest	End of paper	1	-
TOTAL SCORE (% of max possible)			11 (48%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 19: Phillips (1994)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of setting or methodology
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Methods	1	-
5	Setting and location	Methods	0.5	Limited information provided
6	Study perspective	Methods	0.5	Not linked to costs
7	Comparators	Methods	1	-
8	Time horizon	-	0	-
9	Discount rate	Methods	0.5	No justification provided
10	Choice of health outcomes	Methods	0.5	Only partial information provided
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	0.5	Lack of information on resource use
14	Currency, price date, and conversion	-	0	-
15	Choice of model	Methods	0.5	No justification provided
16	Assumptions	Methods	0.5	Only partially described
17	Analytical methods	Methods	1	-
18	Study parameters	Methods	0.5	Only partially described
19	Incremental costs and outcomes	Results	1	-
20	Characterising uncertainty	Results	0.5	Sensitivity analysis is not comprehensive
21	Characterising heterogeneity	Results	1	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	0.5	Generalisability and current knowledge not considered
23	Source of funding	End of paper	0.5	No information on role of funder
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			14 (61%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 20: Scott (2002)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	0.5	Economic evaluation not specified
2	Abstract	Abstract	1	-
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Introduction/methods	1	-
5	Setting and location	Introduction	0.5	Limited information on system
6	Study perspective	Abstract	0.5	Not related to costs
7	Comparators	-	0	-
8	Time horizon	Methods	0.5	Not clearly specified
9	Discount rate	Methods	0.5	Not clearly specified
10	Choice of health outcomes	Methods	1	-
11	Measurement of effectiveness	Methods	0.5	Lack of information on identification and synthesis of effectiveness data
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	1	-
14	Currency, price date, and conversion	Methods	0.5	Currency calculations unclear
15	Choice of model	Methods	0.5	Insufficient description
16	Assumptions	Methods	0.5	Only partially completed
17	Analytical methods	Methods	0.5	Only partially completed
18	Study parameters	Methods	1	-
19	Incremental costs and outcomes	Results	0.5	Not disaggregated
20	Characterising uncertainty	Results	1	-
21	Characterising heterogeneity	-	N/A	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	-	0	-
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			13.5 (61%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 21: Sund (2012)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of perspective or uncertainty
3	Background and objectives	Introduction	0.5	Limited information provided
4	Target population and subgroups	Methods	0.5	Incomplete information provided on base case population
5	Setting and location	Introduction/methods	1	-
6	Study perspective	-	0	-
7	Comparators	-	0	-
8	Time horizon	Methods	1	-
9	Discount rate	Methods	1	-
10	Choice of health outcomes	Introduction/methods	1	-
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	Methods	0.5	Incomplete information provided
13	Estimating resources and costs	Methods	1	-
14	Currency, price date, and conversion	Methods	1	-
15	Choice of model	-	N/A	-
16	Assumptions	-	N/A	-
17	Analytical methods	Methods	0.5	Only partially completed
18	Study parameters	Methods	0.5	No variations considered
19	Incremental costs and outcomes	Results	0.5	Only partially completed
20	Characterising uncertainty	Results	0.5	Sensitivity analysis is not comprehensive
21	Characterising heterogeneity	-	N/A	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	End of paper	1	-
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			14 (67%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 22: Tang (1992)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of perspective or uncertainty
3	Background and objectives	Introduction/discussion	0.5	Limited information provided
4	Target population and subgroups	Methods	0.5	No justification provided
5	Setting and location	Methods	0.5	Limited information provided
6	Study perspective	Methods	1	-
7	Comparators	Methods	1	-
8	Time horizon	-	0	-
9	Discount rate	-	0	-
10	Choice of health outcomes	Methods	1	-
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	-	N/A	-
13	Estimating resources and costs	Methods	0.5	Unit cost data incomplete
14	Currency, price date, and conversion	-	0	-
15	Choice of model	-	N/A	-
16	Assumptions	-	N/A	-
17	Analytical methods	Methods	1	-
18	Study parameters	Methods	0.5	Lack of unit cost data
19	Incremental costs and outcomes	Results	0.5	Limited cost information
20	Characterising uncertainty	Results	0.5	Limited analysis of uncertainty
21	Characterising heterogeneity	-	0	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	Start of paper	0.5	No information on role of funder
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			11.5 (55%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 23: Torrance (1999)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	1	-
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Methods	1	-
5	Setting and location	Methods	0.5	System not well-described
6	Study perspective	Methods	1	-
7	Comparators	Introduction/methods	1	-
8	Time horizon	Methods	1	-
9	Discount rate	Methods	1	-
10	Choice of health outcomes	Methods	1	-
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	Methods	1	-
13	Estimating resources and costs	Methods	1	-
14	Currency, price date, and conversion	Methods	1	-
15	Choice of model	-	N/A	-
16	Assumptions	-	N/A	-
17	Analytical methods	Methods	1	-
18	Study parameters	Methods	0.5	Lacking data on uncertainty
19	Incremental costs and outcomes	Results	1	-
20	Characterising uncertainty	Results	1	-
21	Characterising heterogeneity	Results	1	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	End of paper	1	-
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			20 (91%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 24: Van Oostrom (2010)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	1	-
2	Abstract	Abstract	0.5	No mention of setting or uncertainty
3	Background and objectives	Introduction	1	-
4	Target population and subgroups	Methods	1	-
5	Setting and location	Methods	1	-
6	Study perspective	Methods	1	-
7	Comparators	Introduction	1	-
8	Time horizon	Methods	1	-
9	Discount rate	Methods	1	-
10	Choice of health outcomes	Methods	0.5	No justification provided
11	Measurement of effectiveness	Methods	1	-
12	Measurement and valuation of preference based outcomes	Methods	1	-
13	Estimating resources and costs	Methods	1	-
14	Currency, price date, and conversion	Methods	1	-
15	Choice of model		N/A	-
16	Assumptions		N/A	-
17	Analytical methods	Methods	1	-
18	Study parameters	Methods	0.5	Lack of data on variance
19	Incremental costs and outcomes	Results	1	-
20	Characterising uncertainty	Results	0.5	Only partially characterised
21	Characterising heterogeneity	Results	1	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	1	-
23	Source of funding	End of paper	0.5	No information on role of funder
24	Conflicts of interest	End of paper	1	-
TOTAL SCORE (% of max possible)			19.5 (89%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Study 25: Yianni (2005)

	Item	Where reported?	Score ^a	Notes
1	Title	Title	0.5	Unclear.
2	Abstract	Abstract	0.5	No mention of setting, perspective or uncertainty
3	Background and objectives	Introduction	0.5	Lack of detail on decision context
4	Target population and subgroups	Methods	0.5	Limited information provided
5	Setting and location	Methods	0.5	Limited information provided
6	Study perspective	Methods	0	-
7	Comparators	Methods	1	-
8	Time horizon	Methods	0.5	No justification provided
9	Discount rate	-	0	-
10	Choice of health outcomes	Introduction/methods	0.5	No justification provided
11	Measurement of effectiveness	Introduction/methods	0.5	Lack of detail
12	Measurement and valuation of preference based outcomes	Methods	0.5	Lack of detail
13	Estimating resources and costs	Methods/results	1	-
14	Currency, price date, and conversion	Methods	0.5	Limited information provided
15	Choice of model	-	N/A	-
16	Assumptions	-	N/A	-
17	Analytical methods	Methods	0.5	Lack of detail
18	Study parameters	Methods/results	0.5	Lack of data on variability
19	Incremental costs and outcomes	Results	1	-
20	Characterising uncertainty	-	0	-
21	Characterising heterogeneity	-	0	-
22	Study findings, limitations, generalisability, and current knowledge	Discussion	0.5	Generalisability and current knowledge not considered
23	Source of funding	End of paper	0.5	No information on role of funder
24	Conflicts of interest	-	0	-
TOTAL SCORE (% of max possible)			10 (45%)	

^a 1=fully addressed, 0.5=partially addressed, 0=not addressed

Summary of quality ratings for each paper

First author	Year	Concordance	Score for each checklist item																								Total score	Max score	% [†]
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24			
Andresen	2010	Y	1	½	1	1	0	1	0	½	0	½	½		½	0	½	1	0	½	½	½	0	1	0	0	10.5	23	46%
Beale	2012	Y	½	1	1	0	1	1	1	1	½	1	½	½	½	1	0	1	0	0	½	½	0	1	½	1	15	24	63%
Brisson	2006	Limited	½	1	1	1	½	1	1	1	1	1	½	½	½	½	1	1	1	1	½	1	1	1	½	1	20	24	83%
Burr	2012	N	1	1	1	1	1	1	1	1	1	1	1	1	½	½	1	1	1	1	1	1	1	1	1	1	23	24	96%
Caughey	2002	Y	1	½	1	0	1	0	1	0	½	½	1	0	½	1	½	0	0	1	1	½	0	1	½	0	12.5	24	52%
Chen	1998	N	1	½	1	1	0	0	1	0	0	½	0	0	½	½	½	0	0	½	½	½	1	½	0	0	9.5	24	40%
Driessen	2012	N	1	½	½	½	½	1	1	0	0	1	1		1	½			1	0	½	0	0	1	½	0	11.5	21	55%
Evans	1997	Y	1	½	½	1	0	1	1	1	1	1	½	1	½	½	½	1	1	0	1	1	½	½	½	0	16.5	24	69%
Ginsberg	2001	Y	1	½	1	1	1	1	½	1	½	½	½		1	1	½	½	½	½	½	1		1	½	1	16.5	22	75%
Goodchild	2011	Y	1	½	½	1	½	0	0	½	½	1	½		½	0			½	½	½	½		0	½	0	9	20	45%
Grosse	2005	Y	½	0	1	1	1	0	0	0	1	1	1		½	½			0	0	0	½		½	0	0	8.5	20	43%
Holtgrave	1993	Y	1	1	1	1	1	½	0	½	1	1	1		1	½			½	½	0	1	½	1	0	0	14	21	67%
Johannesson	1992	Y	½	½	1	1	½	0	1	1	½	1	1		1	½	½	½	½	½	1	½	0	1	½	0	14.5	23	63%
Liu	2012	Y	1	1	1	1	1	0	1	½	½	1	1	½	½	½	½	1	½	1	1	1	0	1	½	1	18	24	75%
Loisel	2002	Y	1	½	1	½	0	1	½	½	0	½	1		1	1			1	½	1	½	0	½	0	0	12	21	57%
McCormack	2005	N	1	1	1	½	1	1	1	1	½	½	1	½	1	½	½	1	1	1	1	1	1	1	1	1	21	24	88%
McKay	1991	Y	1	½	1	1	1	0	0	0	0	1	1		½	0	0	½	0	0	1	½		1	0	0	10	22	45%
Moradi-Lakeh	2012	Y	1	1	½	½	0	1	1	0	½	½	½		½	0	½	0	0	½	½	½	0	½	½	1	11	23	48%
Phillips	1994	Y	1	½	1	1	½	½	1	0	½	½	1		½	0	½	½	1	½	1	½	1	½	½	0	14	23	61%
Scott	2002	Y	½	1	1	1	½	½	0	½	½	1	½		1	½	½	½	½	1	½	1		1	0	0	13.5	22	61%
Sund	2012	Y	1	½	½	½	1	0	0	1	1	1	1	½	1	1			½	½	½	½		1	1	0	14	21	67%
Tang	1998	Y	1	½	½	½	½	1	1	0	0	1	1		½	0			1	½	½	½	0	1	½	0	11.5	21	55%
Torrance	1999	Limited	1	1	1	1	½	1	1	1	1	1	1	1	1	1			1	½	1	1	1	1	1	0	20	22	91%
Van Oostrom	2010	Y	1	½	1	1	1	1	1	1	1	½	1	1	1	1			1	½	1	½	1	1	½	1	19.5	22	89%
Yianni	2005	Limited	½	½	½	½	½	0	1	½	0	½	½	½	1	½			½	½	1	0	0	½	½	0	10	22	45%
% studies meeting criteria			88%	66%	86%	78%	62%	58%	68%	54%	52%	80%	78%	58%	72%	52%	50%	63%	56%	52%	70%	64%	42%	82%	44%	32%	Mean score		63%
Ranking of item			1	10	2	5	13	15	9	17	18	4	5	14	7	18	21	12	16	18	8	11	23	3	22	24			
																				Average score for studies published in 2002 or earlier							60%		
																				Average score for studies published post-2002							66%		
																				Average score for studies showing concordance							60%		
																				Average score for studies not showing concordance							71%		

[†] Studies classified as low quality (fulfilled <50% of applicable quality criteria), average quality (fulfilled between 50% and 80% of applicable criteria) or high quality (fulfilled >80% of applicable criteria)

Appendix 3: References included in the literature review which considers the impact of choice of economic evaluation approach on adoption decisions

1. Andresen MA, Boyd N (2010) A cost-benefit and cost-effectiveness analysis of Vancouver's supervised injection facility. *International Journal of Drug Policy* 21: 70-76.
2. Beale SJ, Bending MW, Trueman P, Naidoo B (2012) Should we invest in environmental interventions to encourage physical activity in England? An economic appraisal. *European Journal of Public Health* 22: 869-873.
3. Brisson M, Edmunds WJ (2006) Impact of model, methodological, and parameter uncertainty in the economic analysis of vaccination programs. *Medical Decision Making* 26: 434-446.
4. Burr J, Botello Pinzon P, Takwoingi Y, Hernandez R, Vazquez-Montes M, et al. (2012) Surveillance for ocular hypertension: an evidence synthesis and economic evaluation. *Health Technology Assessment* 16: 271.
5. Caughey AB, Kuppermann M, Norton ME, Washington AE (2002) Nuchal translucency and first trimester biochemical markers for Down syndrome screening: A cost-effectiveness analysis. *American Journal of Obstetrics and Gynecology* 187: 1239-1245.
6. Chen S, Shaheen A, Garber A (1998) Cost-effectiveness and cost-benefit analysis of using methotrexate vs Goeckerman therapy for psoriasis: A pilot study. *Archives of Dermatology* 134: 1602-1608.
7. Driessen M, Bosmans J, Proper K, Anema J, Bongers P, et al. (2012) The economic evaluation of a Participatory Ergonomics programme to prevent low back and neck pain. *Work* 41 Suppl 1: 2315-2320.
8. Evans KW (1997) Economic evaluation of oral sumatriptan compared with oral caffeine/ergotamine for migraine. *Pharmacoeconomics* 12: 565-577.
9. Ginsberg GM, Slater PE, Shouval D (2001) Cost-benefit analysis of a nationwide infant immunization programme against hepatitis A in an area of intermediate endemicity. *Journal of Hepatology* 34: 92-99.
10. Goodchild M, Sahu S, Wares F, Dewan P, Shukla RS, et al. (2011) A cost-benefit analysis of scaling up tuberculosis control in India. *International Journal of Tuberculosis and Lung Disease* 15: 358-362.
11. Grosse SD, Waitzman NJ, Romano PS, Mulinare J (2005) Reevaluating the benefits of folic acid fortification in the United States: economic analysis, regulation, and public health. *American Journal of Public Health* 95: 1917-1922.
12. Holtgrave DR, Valdiserri RO, Gerber AR, Hinman AR (1993) Human immunodeficiency virus counseling, testing, referral, and partner notification services: A cost-benefit analysis. *Archives of Internal Medicine* 153: 1225-1230.
13. Johannesson M, Fagerberg B (1992) A health-economic comparison of diet and drug treatment in obese men with mild hypertension. *Journal of Hypertension* 10: 1063-1070.
14. Liu Y, Mack KA, Diekman ST (2012) Smoke alarm giveaway and installation programs: an economic evaluation. *American Journal of Preventive Medicine* 43: 385-391.
15. Loisel P, Lemaire J, Poitras S, Durand M-J, Champagne F, et al. (2002) Cost-benefit and cost-effectiveness analysis of a disability prevention model for back pain management: a six year follow up study. *Occupational and Environmental Medicine* 59: 807-815.
16. McCormack K, Wake B, Perez J, Fraser C, Cook J, et al. (2005) Laparoscopic surgery for inguinal hernia repair: systematic review of effectiveness and economic evaluation. *Health Technology Assessment* 9: 218.

17. McKay NL, Phillips KM (1991) An economic evaluation of mandatory premarital testing for HIV. *Inquiry* 28: 236-248.
18. Moradi-Lakeh M, Shakerian S, Esteghamati A (2012) Immunization against Haemophilus Influenzae type B in Iran; cost-utility and cost-benefit analyses. *International Journal of Preventive Medicine* 3: 332-340.
19. Phillips KA, Lowe RA, Kahn JG, Lurie P, Avins AL, et al. (1994) The cost-effectiveness of HIV testing of physicians and dentists in the United States. *JAMA: The Journal of the American Medical Association* 271: 851-858.
20. Scott RD, Meltzer MI, Erickson LJ, De Wals P, Rosenstein NE (2002) Vaccinating first-year college students living in dormitories for Meningococcal disease: an economic analysis. *American Journal of Preventive Medicine* 23: 98-105.
21. Sund B, Svensson L, Rosenqvist M, Hollenberg J (2012) Favourable cost-benefit in an early defibrillation programme using dual dispatch of ambulance and fire services in out-of-hospital cardiac arrest. *European Journal of Health Economics* 13: 811-818.
22. Tang J, Wang B, White PF, Watcha MF, Qi J, et al. (1998) The effect of timing of ondansetron administration on its efficacy, cost-effectiveness, and cost-benefit as a prophylactic antiemetic in the ambulatory setting. *Anesthesia & Analgesia* 86: 274-282.
23. Torrance G, Walker V, Grossman R, Mukherjee J, Vaughan D, et al. (1999) Economic evaluation of ciprofloxacin compared with usual antibacterial care for the treatment of acute exacerbations of chronic bronchitis in patients followed for 1 year. *Pharmacoeconomics* 16: 499-520.
24. van Oostrom SH, Heymans MW, de Vet HCW, van Tulder MW, van Mechelen W, et al. (2010) Economic evaluation of a workplace intervention for sick-listed employees with distress. *Occupational and Environmental Medicine* 67: 603-610.
25. Yianni J, Green AL, McIntosh E, Bittar RG, Joint C, et al. (2005) The costs and benefits of deep brain stimulation surgery for patients with dystonia: An initial exploration. *Neuromodulation* 8: 155-161.

Appendix 4: Summary of references reporting the results of both a welfarist and an extra-welfarist economic evaluation in the same study

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Andresen (2010)	Not indicated	Addiction treatment (Canada) *	Introduction of a supervised injection facility versus do nothing	CEA	N/A	No explicit justification provided	Expected value of a prevented death calculated	CEA: ICER = CDN\$42,857. Outcome measure not clearly specified. Acceptable upper limit for ICER not specified CBA: Benefit-cost ratio = 5:1 (Estimated benefit CDN\$7.56m, estimated cost CDN\$1.5m)	Yes Results not explicitly compared, but the cost-benefit ratios are positive, and the average ICER is within a commonly used acceptable range (<\$100,000)	No
Beale (2012)	Not commercially funded	Promoting physical activity (UK) *	Four different cycling interventions compared, each versus a do-nothing alternative	CUA	EQ-5D	Yes Note that adopting a CUA approach alone in this context may underestimate intervention benefits, and a broader approach may be required	Unclear (ranges for monetary benefits devised based on literature review, but unclear how these estimates were combined)	CUA: ICERs ranged from £100-£10,000/QALY, depending on the intervention CBA: Benefit-cost ratio = 11:1 (standardised across all interventions)	Yes Results not explicitly compared, but the authors note that their findings across the different approaches present a consistent case for adoption	No

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Brisson (2006)	Not commercially funded	Varicella vaccination (UK) *	Routine varicella vaccination versus no vaccination	CUA	HUI	Yes Note that QALYs may not capture all of the benefits of vaccination, as compared to WTP/CBA	Contingent Valuation	CEA: Point estimate for ICER not provided, but it is noted that virtually all of the ICERs generated by PSA fall below a threshold level of £30,000/QALY in the baseline analysis from a NHS perspective. When a societal perspective is taken, the CEA results are more variable CBA: The likelihood of varicella vaccination resulting in a positive net benefit is estimated to be 100%, regardless of perspective	Limited The authors compare the two forms of economic evaluation by considering whether results fell above or below their respective thresholds, and conclude that the economic desirability of varicella vaccination is uncertain using CUA, but highly beneficial using CBA	No

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Burr (2012)	Not commercially funded	Surveillance for ocular hypertension (UK)	Five surveillance pathways compared with different monitoring intervals and treatment thresholds	CEA, CUA	EQ-5D	No explicit justification provided	DCE	CEA: ICERs varied from £3-415 per measure of effectiveness. Three measures were used for different degrees of progression to open-angle glaucoma (OAG). An intensive surveillance strategy ("NICE Intensive") is likely to be the most-cost-effective approach CUA: The most cost-effective strategy is a less intensive hospital led surveillance approach ("SOH Hospital") at £85,213/QALY gained, however this ICER is well above the usual NICE threshold. "Treat all" is therefore the optimal approach CBA: The only strategy with a positive net benefit is the SOH Hospital approach	No Each approach suggests a different adoption decision	<u>No</u>

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Caughey (2002)	Not commercially funded	Prenatal screening (USA)	Nuchal translucency (NT) versus first-trimester serum screening versus combined NT/first-trimester serum screening vs. second-trimester α -fetoprotein (XAFP) screening (baseline)	CEA, CUA	SG	No explicit justification provided	Expected value of a prevented case calculated	CEA: First-trimester vs. baseline = \$160,266/case identified NT vs. first-trimester = \$98,381/case identified Combined vs. NT = \$319,934/case identified Threshold of \$577,248/case identified used (societal costs of Down syndrome foetus) CBA: Benefit-cost ratios: First-trimester = 5:1 NT = 5:1 Combined = 2:1 CUA: NT vs. baseline = \$86,566/QALY gained Combined vs. baseline = \$76,552/QALY gained First-trimester vs. baseline = Dominated Threshold of \$100,000/QALY used CUA results were less favourable (>\$100,000/QALY if a societal perspective was taken)	Yes The three forms of economic evaluation were compared in terms of whether results fell above or below their respective thresholds and they were generally below. However the CEA threshold differs significantly from the CUA threshold, which is not a common approach	<u>No</u>

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Chen (1998)	Not indicated	Psoriasis (USA)	Methotrexate (liquid or tablet) vs. Goeckerman therapy (GKM)	CUA	VAS	Yes WTP may be a more sensitive method for valuing outcomes in non-life-threatening disorders	WTP questionnaire	CUA: Compared to no therapy, all three treatments were always cost-effective (using a threshold of \$35,000). GKM was in some instances more cost-effective than liquid methotrexate but this depended on whose utility estimates were used and the severity of psoriasis CBA: Generally all three treatments are not cost beneficial compared to no therapy, with the greatest net benefit observed for liquid methotrexate. GKM provided no net benefit when compared with liquid methotrexate. There is some variability in the results depending on whose WTP estimates are used	No Although all three treatments were always cost-effective vs. no therapy, they were rarely cost-beneficial	No However the authors focus on issues surrounding the implementation and use of CUA rather than CBA

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Driessen (2012)	Not commercially funded	Managing back pain (Netherlands)	Participatory Ergonomics (PE) programme versus no programme	CEA	N/A	No explicit justification provided	FCA	CEA: PE programme cost €23,749 for an additional unit of effectiveness CBA: Loss of €78 per worker	No Study concluded that there was concordance between the CEA and CBA results. However the authors misinterpreted the CEA results as indicating that the intervention was not cost-effective. It is, however, likely that the ICER would fall under most CEA thresholds.	<u>No</u>

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Evans (1997)	Not commercially funded	Migraine (Canada)	Oral sumatriptan with oral caffeine/ergotamine for the treatment of migraine attacks	CEA, CUA	Multi-attribute utility instrument	No explicit justification provided	HCA	CEA: Cost per attack averted from the health service (societal) perspective: CDN\$98 (Dominant) CUA: Cost per QALY gained from the health service (societal) perspective: CDN\$29,366 (Dominant) CBA: Net benefit = CDN\$42/year	Yes The authors note that the economic evaluations are consistent with each other: they indicate a positive net economic benefit from the societal perspective, and improved health outcomes achieved at moderately acceptable incremental costs from the health service perspective.	No

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Ginsberg (2001)	Not commercially funded	Hepatitis A (Israel) *	Introduction of a nationwide active immunisation programme versus current practice	CEA	N/A	No explicit justification provided	GNP value of a life-year lost	CEA: US\$314/case averted CBA: Benefit-cost ratio = 3:1	Yes The two forms of economic evaluation were not explicitly compared but both indicated that the introduction of a nationwide active immunisation programme would provide a net economic benefit to society. The average ICER is well within a commonly used acceptable range (<\$100,000)	<u>No</u>
Goodchild (2011)	Not commercially funded	TB (India) *	Scaled up TB control versus previous limited approach	CUA	DALYs	No explicit justification provided	GNP value of a statistical life	CUA: US\$26/DALY gained CBA: Benefit-cost ratio = 115:1	Yes The two forms of economic evaluation were not explicitly compared, but the authors judged the intervention to be financially worthwhile in both cases	<u>No</u>

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Grosse (2005)	Not commercially funded	Folic acid fortification (USA) *	Ex-post economic evaluation of the introduction of folic acid fortification in the USA	CEA	N/A	No explicit justification provided However, noted that US guidance calls for both CEA and CBA when conducting an economic evaluation of a public health intervention	Expected value of a prevented death calculated	CEA: \$4,902/neural tube defect averted CBA: Benefit-cost ratio = 142:1	Yes Results not explicitly compared, but the cost-benefit ratio is positive, and the average ICER is within a commonly used acceptable range (<\$100,000)	No
Holtgrave (1993)	Not commercially funded	HIV (USA) *	Provision of HIV prevention services versus no prevention services in the USA	CEA	N/A	No explicit justification provided	HCA	CEA: \$25,780/infection averted CBA: Benefit-cost ratio = 20:1	Yes The two forms of economic evaluation were not explicitly compared but both indicated that the provision of HIV prevention services would provide a net economic benefit to society. The average ICER is well within a commonly used acceptable range (<\$100,000)	No

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Johannesson (1992)	Not commercially funded	Hypertension (Sweden)	Use of diet therapy versus use of drug therapy (baseline)	CEA	N/A	No explicit justification provided	WTP questionnaire	CEA: Five scenarios considered. In three of these, drug treatment dominated. The final two scenarios favoured diet therapy at SEK 46,000 (£4,311) and SEK 205,000 (£19,211) CBA: Both treatments resulted in a loss compared with no treatment, and the difference between them was negligible	Yes Both economic evaluations present a mixed picture. The authors refer to a threshold value when considering the results of the CEA but do not specify what this is	Yes The authors note that CBA is theoretically better able to handle such problems, as patients take into account their costs of side-effects, leisure time etc. when stating their WTP
Liu (2012)	Not commercially funded	Smoke alarm use (USA) *	Evaluation of programmes to either giveaway or install smoke alarms, versus do nothing	CEA, CUA	EQ-5D	No explicit justification provided	Value of a statistical life	CEA: \$46,673 per life-year gained (giveaway programme) / \$38,285 per life-year gained (installation programme) CUA: \$51,404 per QALY gained (giveaway programme) / \$42,165 per QALY gained (installation programme) CBA: \$48,827 (giveaway programme) / \$137,119 (installation programme)	Yes The results of the CEA and CUA are likely to fall under commonly used thresholds, and the CBA indicates a net benefit	Yes The authors note that the CBA results may be unreliable (concerns over the methods used to assign a monetary value to life and health-related quality of life), so focus on CEA/CUA results

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Loisel (2002)	Not commercially funded	Managing back pain (Canada) *	Standard care versus Clinical rehabilitation versus Occupational intervention versus Combined clinical rehabilitation and occupational intervention (Sherbrooke model)	CEA	N/A	No explicit justification provided	HCA (although somewhat unclear)	CEA: All three interventions were cost saving and improved effectiveness in terms of days on full benefits averted. The Sherbrooke model was the most effective CBA: All three interventions were cheaper and increased the monetary value of benefits as compared to standard care. There was no statistical difference between the results in the three intervention arms	Yes Both economic evaluations indicate that the three interventions are financially worthwhile as compared to usual practice, and in both cases it is unclear which of the three interventions should be preferred	No

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
McCormack (2005)	Not commercially funded	Hernia surgery (UK)	Laparoscopic surgery versus open mesh surgery for inguinal hernia repair	CUA	EQ-5D	Yes Uncertainty concerning the extent to which EQ-5D valuations based on population data match valuations based on the preferences of patients	DCE	CUA: Results are dependent on the time horizon adopted and the assumptions made concerning operation length and length of stay. Laparoscopic surgery is the most cost-effective option when longer time horizons are considered. CBA: The results of the DCE were integrated into the economic model and indicated that laparoscopic surgery was unlikely to be associated with net benefits compared to open mesh methods	No The CUA suggests that laparoscopic surgery should be adopted, while the CBA suggests that open mesh methods provide a greater net benefit	<u>No</u>

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
McKay (1991)	Not indicated	HIV testing (USA) *	Introduction of a mandatory premarital testing program for HIV versus current practice (no program)	CEA	N/A	No explicit justification provided	HCA and WTP estimate taken from literature	CEA: Low transmission rates = \$127,000/case prevented; High transmission rates = \$70,000/case prevented CBA: Benefit-cost ratio for low transmission rates = 16:1; high transmission rates = 28:1	Yes The authors note that their study provides a positive assessment of premarital testing for HIV, but do not explicitly compare the results of the two forms of economic evaluation	No
Moradi-Lakeh (2012)	Not commercially funded	Hib vaccination (Iran) *	Introduction of a Hib vaccination programme versus current practice (do nothing)	CEA, CUA	DALYs	No explicit justification provided	HCA	CEA: US\$6,754/case averted; US\$136,300/death averted CUA: US\$4,535/DALY averted CBA: Benefit-cost ratio = 2:1	Yes Although the cost per death averted is slightly higher than a commonly used acceptable range (<\$100,000) there is general concordance between the results, which indicate that the intervention provides value for money	No

Phillips (1994)	Not commercially funded	HIV testing (USA) *	Mandatory testing or increased voluntary testing of surgeons and dentists versus do nothing	CEA	N/A	No explicit justification provided	Value of a statistical life calculated	CEA: Surgeons – ICERs range from \$29m per patient infection averted to cost saving, depending on prevalence & transmission risk Dentists – ICERs range from \$447m per patient infection averted to cost saving, depending on prevalence and transmission risk CBA: An explicit benefit-cost ratio is not calculated due to uncertainty over the valuation of morbidity/mortality. However the authors calculate what the value of one patient infection averted would have to exceed for the benefits of testing to exceed the costs - Surgeons – restrictions placed on practice = \$20m; prohibited from practicing = \$43m - Dentists – restrictions placed on practice = \$13m; prohibited from practicing = \$28m	Yes Neither form of economic evaluation provides a definitive answer, and both forms produce highly variable results	No

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Scott (2002)	Not commercially funded	Vaccination for meningococcal disease (USA) *	Vaccination for meningococcal disease versus no vaccination	CEA	N/A	No explicit justification provided	HCA and Value of a statistical life calculated	CEA: Cost per case averted ranged from \$0.6m to \$1.9m. Cost per death averted ranged from \$7m to \$20m. Cost per life-year saved ranged from \$62,042 to \$489,185 CBA: Net benefit ranged from -\$11m to -\$49m	Yes There was generally concordance between results – both types of economic evaluation indicated that the vaccination programme was unlikely to provide value for money, except in a very specific best-case scenario	<u>No</u>
Sund (2012)	Not commercially funded	Heart disease (Sweden)	Dual dispatch of ambulance and fire services in out-of-hospital cardiac arrest versus do nothing	CEA, CUA	HUI	No explicit justification provided	Value of a statistical life calculated	CEA: €60,000 per life saved CUA: €13,000 per QALY gained CBA: Benefit-cost ratio = 36:1	Yes The results of the CEA and CUA are likely to fall under commonly used thresholds, and the CBA indicates a net benefit	<u>No</u>

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Tang (1998)	Commercially funded	Antiemetic treatment (USA)	Administration of antiemetic before surgery versus administration after surgery versus split dose before and after versus placebo	CEA	N/A	No explicit justification provided	WTP questionnaire	CEA: The average cost-effectiveness ratios (cost per patient with a complete response) were: Placebo \$53; Split dose \$64; Before \$46, and After \$36 CBA: Benefit-cost ratio = 5:1	Yes The benefit-cost ratio is favourable and the ACERs are likely to fall under any reasonable threshold for cost-effectiveness, however incremental analysis was not performed	No

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Torrance (1999)	Commercially funded	Bronchitis (Canada)	Ciprofloxacin versus usual antibacterial care for the treatment of chronic bronchitis	CEA, CUA	HUI	No explicit justification provided	WTP questionnaire	CEA: Ministry of health (societal) perspective: \$304 (\$209)/symptom day averted CUA: Ministry of health (societal) perspective: \$27,025 (\$18,588)/QALY gained CBA: Net benefit from Ministry of health (societal) perspective using mean values: - \$1207 (-\$945) For the Ministry of health (societal) perspective using median values, exact figures were not given, but the authors stated that the net benefit remained negative	Limited The authors compared the three economic evaluations and concluded that the CEA and CUA provided evidence of value for money, with the CUA showing strong evidence for adoption from the societal perspective (moderate evidence from the ministry of health perspective. However the CBA results are not favourable	Yes The authors favour CEA and CUA in this context, noting that they do not have much faith in the CBA results

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Van Oostrom (2010)	Not commercially funded	Mental health (Netherlands)	Workplace intervention for sick-listed employees with distress versus Usual care	CEA, CUA	EQ-5D	No explicit justification provided	HCA and FCA	CEA: €627/day reduction in sick leave. This estimate is not robust because the differences in costs and effects spans zero CUA: Workplace intervention dominated by usual care CBA: Net monetary benefit ranged from €1700 to €1987 but these estimates are not robust as the confidence intervals span zero	Yes The results of the economic evaluations are either not robust (CEA, CBA) or indicate that the workplace intervention is dominated by usual care	<u>No</u>

Reference (publication date)	Source of funding	Intervention context (study setting) ^a	Comparison	Type of extra-welfarist economic evaluation	Method used to estimate utilities within CUA	Justification provided for welfarist approach?	Method used to monetise benefits within welfarist approach	Economic evaluation results	Concordance between results?	Most appropriate form of economic evaluation identified?
Yianni (2005)	Not commercially funded	Neurosurgery (UK)	Deep brain stimulation (DBS) versus no intervention	CUA	EQ-5D	Yes The authors note that CBA is arguably the type of health economic evaluation that is best suited to decide whether a healthcare intervention is worth implementing	WTP questionnaire	CUA: £33,980/QALY gained CBA: Using mean (median) WTP, Benefits-Costs = £259,289 (-£11,942)	Limited The CUA provides weak evidence for cost-effectiveness: the ICER falls at the upper end of the acceptable range in the UK. The CBA provides some evidence of value for money, but this depends on the definition of WTP, which has a significant impact on the results	No

ACER = average cost-effectiveness ratio; CBA = cost-benefit analysis; CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DALY = disability-adjusted life years; DCE = discrete choice experiment; EQ-5D = EuroQol five dimensions questionnaire; FCA = friction cost approach; GNP = gross national product; HCA = human capital approach; HUI = health utilities index questionnaire; ICER = incremental cost-effectiveness ratio; NICE = National Institute for Health and Care Excellence; QALY = quality-adjusted life year; SG = standard gamble; VAS = visual analogue scale; WTP = willingness-to-pay. ^a Public health interventions are marked with a *.

Appendix 5: Search strategies used to identify references which summarise the methodological challenges faced when conducting economic evaluations of genomic interventions

Embase (1988 to present)

genetic*.ti,ab.
exp genetics/
genomic*.ti,ab.
exp genomics/
pharmacogenetic*.ti,ab.
exp pharmacogenetics/
pharmacogenomic*.ti,ab.
exp pharmacogenomics
or/1-8
methodol*.ti,ab.
case stud*.ti,ab.
method*.ti.
or/10-12
economic adj3 evaluation. ti,ab.
exp health economics/ or exp economic evaluation/
cost adj3 effectiveness. ti,ab.
exp "cost effectiveness analysis"/
cost adj3 benefit. ti,ab.
exp "cost benefit analysis"/
cost adj3 utility. ti,ab.
exp "cost utility analysis"/ or exp "quality of life"/
cost adj3 minimi#ation. ti,ab.
exp "cost minimization analysis"
cost adj3 consequence.ti,ab.
cost*. ti,ab.
exp "health care cost"/ or exp "cost"/
pharmacoeconomic*.ti,ab.
life year*.ti,ab.
quality adjusted.ti,ab.
utilit*.ti,ab.
benefit.ti,ab.
Or/14-31
9 and 13 and 32
limit 33 to english language
editorial.pt
letter.pt
note.pt
(rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep or primate or primates).ti,ab.
or/35-38
34 not 39

limit 40 to human
limit 41 to yr="1990 -Current"

Medline In-Process & Other Non-Indexed Citations and Medline 1946 to present + Medline Daily Update

genetic*.ti,ab.
exp genetics/
genomic*.ti,ab.
exp genomics/
pharmacogenetic*.ti,ab.
exp pharmacogenetics/
pharmacogenomic*.ti,ab.
exp pharmacogenomics
or/1-8
methodol*.ti,ab.
case stud*.ti,ab.
method*.ti.
or/10-12
economic adj3 evaluation.ti,ab.
exp Economics, Medical/ or exp Health Care Rationing/
cost adj3 effectiveness. ti,ab.
cost adj3 benefit.ti,ab.
exp cost benefit analysis/
cost adj3 utility.ti,ab.
exp "quality of life"/
cost adj3 minimi#ation.ti,ab.
cost adj3 consequence.ti,ab.
cost*. ti,ab.
exp "costs and cost analysis"/ or exp Health Care Costs/
pharmacoeconomic*.ti,ab.
life year*.ti,ab.
quality adjusted.ti,ab.
utilit*.ti,ab.
benefit.ti,ab.
exp Quality-Adjusted Life Years/
or/14-29
9 and 13 and 30
limit 31 to english language
editorial.pt
letter.pt
(rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep or primate or primates).ti,ab.
or/33-35
32 not 36
limit 37 to human
limit 39 to yr="1990 -Current"

Econlit

(ti,ab(genetic*) OR ti,ab(genomic*) OR ti,ab(pharmacogenetic*) OR ti,ab(pharmacogenomic*))
AND (ti,ab(methodol*) OR ti,ab("case stud*") OR ti,ab(method*)) AND (ti,ab(economic NEAR/3
evaluation) OR ti,ab(cost NEAR/3 effectiveness) OR ti,ab(cost NEAR/3 benefit) OR ti,ab(cost
NEAR/3 utility) OR ti,ab(cost NEAR/3 minimi?ation) OR ti,ab(cost*) OR ti,ab(pharmacoeconomic*)
OR ti,ab("life year") OR ti,ab("quality adjusted") OR ti,ab(utilit*) OR ti,ab(benefit))

Limited to English language, excluding dissertations, from 1990 until 16/04/13

The Cochrane Library

Searched for the following in title, abstract and keywords (combined with AND):

genetic* or genomic* or pharmacogenetic* or pharmacogenomic*
method* or case stud*
economic evaluation or cost benefit or cost utility or cost effectiveness or cost minimisation or cost
consequence or cost* or pharmacoeconomic* or life year* or quality adjusted* or utilit* or benefit

Centre for Reviews and Dissemination

Searched for the following across all 3 databases (combined with AND):

genetic* or genomic* or pharmacogenetic* or pharmacogenomic* [any field]
method* or case stud* [title]
economic evaluation or cost benefit or cost utility or cost effectiveness or cost minimisation or cost
consequence or cost* or pharmacoeconomic* or life year* or quality adjusted* or utilit* or benefit
[any field]

Date limited to 1990/01/01-2013/04/16

Appendix 6: References included in the literature review which summarises the methodological challenges faced when conducting economic evaluations of genomic interventions

1. Asch DA, Hershey JC, Pauly MV, Patton JP, Kathryn Jedrzejewski M, Mennuti MT: Genetic screening for reproductive planning: Methodological and conceptual issues in policy analysis. *American Journal of Public Health* 86(5), 684-690 (1996).
2. Bajaj PS, Veenstra DL: A risk-benefit analysis of factor V Leiden testing to improve pregnancy outcomes: a case study of the capabilities of decision modeling in genomics. *Genetics in Medicine*, (2012).
3. Beauchamp JP, et al.: Molecular Genetics and Economics. *Journal of Economic Perspectives* 25(4), 57-82 (2011).
4. Caughey AB: Cost-effectiveness analysis of prenatal diagnosis: Methodological issues and concerns. *Gynecologic and Obstetric Investigation* 60(1), 11-18 (2005).
5. Cohen JS, Biesecker BB: Quality of life in rare genetic conditions: A systematic review of the literature. *American Journal of Medical Genetics, Part A* 152(5), 1136-1156 (2010).
6. Conti R, Veenstra DL, Armstrong K, Lesko LJ, Grosse SD: Personalized Medicine and Genomics: Challenges and Opportunities in Assessing Effectiveness, Cost-Effectiveness, and Future Research Priorities. *Medical Decision Making* 30(3), 328-340 (2010).
7. De Leon J: Pharmacogenomics: The promise of personalized medicine for CNS disorders. *Neuropsychopharmacology* 34(1), 159-172 (2009).
8. Deverka PA, Schully SD, Ishibe N et al.: Stakeholder assessment of the evidence for cancer genomic tests: Insights from three case studies. *Genetics in Medicine* 14(7), 656-662 (2012).
9. Djalalov S, Musa Z, Mendelson M, Siminovitch K, Hoch J: A review of economic evaluations of genetic testing services and interventions (2004-2009). *Genetics in Medicine* 13(2), 89-94 (2011).
10. Faulkner E, Annemans L, Garrison L et al.: Challenges in the Development and Reimbursement of Personalized Medicine—Payer and Manufacturer Perspectives and Implications for Health Economics and Outcomes Research: A Report of the ISPOR Personalized Medicine Special Interest Group. *Value in Health* 15(8), 1162-1171 (2012).
11. Feero W, Wicklund C, Veenstra DL: The economics of genomic medicine: Insights from the iom roundtable on translating genomic-based research for health. *JAMA* 309(12), 1235-1236 (2013).
12. Foster MW, Mulvihill JJ, Sharp RR: Evaluating the utility of personal genomic information. *Genetics in Medicine* 11(8), 570-574 (2009).
13. Garau M, Towse A, Garrison L, Housman L, Ossa D: Can and should value based pricing be applied to molecular diagnostics? OHE research paper 12/03, (2013).
14. Giacomini M, Miller F, O'Brien BJ: Economic considerations for health insurance coverage of emerging genetic tests. *Community Genetics* 6(2), 61-73 (2003).
15. Goddard KAB, Knaus WA, Whitlock E et al.: Building the evidence base for decision making in cancer genomic medicine using comparative effectiveness research. *Genetics in Medicine*, (2012).
16. Grosse SD, Kalman L, Khoury MJ: Evaluation of the validity and utility of genetic testing for rare diseases. *Advances in Experimental Medicine and Biology* 686(pp 115-131), (2010).
17. Grosse SD, Khoury MJ: What is the clinical utility of genetic testing? *Genetics in Medicine* 8(7), 448-450 (2006).
18. Grosse SD, McBride CM, Evans JP, Khoury MJ: Personal utility and genomic information: Look before you leap. *Genetics in Medicine* 11(8), 575-576 (2009).

19. Grosse SD, Rogowski WH, Ross LF, Cornel MC, Dondorp WJ, Khoury MJ: Population screening for genetic disorders in the 21st century: Evidence, economics, and ethics. *Public Health Genomics* 13(2), 106-115 (2010).
20. Grosse SD, Wordsworth S, Payne K: Economic methods for valuing the outcomes of genetic testing: beyond cost-effectiveness analysis. *Genetics in Medicine* 10(9), 648-654 (2008).
21. Gurwitz D, Zika E, Hopkins MM, Gaisser S, Ibarreta D: Pharmacogenetics in Europe: barriers and opportunities. *Public Health Genomics* 12(3), 134-141 (2009).
22. Guzauskas GF, Garrison LP, Stock J, Au S, Doyle DL, Veenstra DL: Stakeholder perspectives on decision-analytic modeling frameworks to assess genetic services policy. *Genetics in Medicine* 15(1), 84-87 (2013).
23. Hall J, Viney R, Haas M: Taking a count: the evaluation of genetic testing. *Aust N Z J Public Health* 22(7), 754-758 (1998).
24. Hall PS, McCabe C, Stein RC, Cameron D: Economic Evaluation of Genomic Test-Directed Chemotherapy for Early-Stage Lymph Node-Positive Breast Cancer. *Journal of the National Cancer Institute* 104(1), 56-66 (2012).
25. Issa AM: Evaluating the value of genomic diagnostics: Implications for clinical practice and public policy. *Advances in Health Economics and Health Services Research* 19(pp 191-206), (2008).
26. Jarrett J, Mugford M: Genetic Health Technology and Economic Evaluation: A Critical Review. *Applied Health Economics and Health Policy* 5(1), 27-35 (2006).
27. Khoury MJ, Gwinn M, Dotson WD, Bowen MS: Is there a need for PGxceptionalism? *Genetics in Medicine* 13(10), 866-867 (2011).
28. Khoury MJ, Rich EC, Randhawa G, Teutsch SM, Niederhuber J: Comparative effectiveness research and genomic medicine: An evolving partnership for 21st century medicine. *Genetics in Medicine* 11(10), 707-711 (2009).
29. Ladabaum U, Wang G, Terdiman J et al.: Strategies to identify the Lynch syndrome among patients with colorectal cancer. *Annals of Internal Medicine* (2), 69-79 (2011).
30. Lin X, Tang W, Ahmad S et al.: Applications of targeted gene capture and next-generation sequencing technologies in studies of human deafness and other genetic disabilities. *Hearing Research* 288(1-2), 67-76 (2012).
31. Malhotra AK, Zhang JP, Lencz T: Pharmacogenetics in psychiatry: Translating research into clinical practice. *Molecular Psychiatry* 17(8), 760-769 (2012).
32. Payne K: Towards an economic evidence base for pharmacogenetics: consideration of outcomes is key. *Pharmacogenomics* 9(1), 1-4 (2007).
33. Payne K: Fish and chips all round? Regulation of DNA-based genetic diagnostics. *Health Economics* 18(11), 1233-1236 (2009).
34. Payne K, McAllister M, Davies LM: Valuing the economic benefits of complex interventions: when maximising health is not sufficient. *Health Economics* 22(3), 258-271 (2012).
35. PHG Foundation: Next Steps In The Sequence: The Implications Of Whole Genome Sequencing For Health In The UK. (2011).
36. Phillips KA, Liang SY, Van Bebber S et al.: Challenges to the translation of genomic information into clinical practice and health policy: Utilization, preferences and economic value. *Current Opinion in Molecular Therapeutics* 10(3), 260-266 (2008).
37. Ransohoff DF, Khoury MJ: Personal genomics: information can be harmful. *European Journal of Clinical Investigation* 40(1), 64-68 (2010).
38. Regier DA, Friedman JM, Marra CA: Value for money? Array genomic hybridization for diagnostic testing for genetic causes of intellectual disability. *American Journal of Human Genetics* (5), 765-772 (2010).
39. Rogowski W: Genetic screening by DNA technology: A systematic review of health economic evidence. *International Journal of Technology Assessment in Health Care* 22(3), 327-337 (2006).

40. Rogowski W: Current impact of gene technology on healthcare. A map of economic assessments. *Health Policy* 80(2), 340-357 (2007).
41. Rogowski WH: The cost-effectiveness of screening for hereditary hemochromatosis in Germany: a remodeling study. *Medical Decision Making* (2), 224-238 (2009).
42. Rogowski WH, Grosse SD, John J et al.: Points to consider in assessing and appraising predictive genetic tests. *Journal of Community Genetics* 1(4), 185-194 (2010).
43. Rogowski WH, Grosse SD, Khoury MJ: Challenges of translating genetic tests into clinical and public health practice. *Nature Reviews Genetics* 10(7), 489-495 (2009).
44. Roth JA, Garrison Jr LP, Burke W, Ramsey SD, Carlson R, Veenstra DL: Stakeholder perspectives on a risk-benefit framework for genetic testing. *Public Health Genomics* 14(2), 59-67 (2011).
45. Sanderson S, Zimmern R, Kroese M, Higgins J, Patch C, Emery J: How can the evaluation of genetic tests be enhanced? Lessons learned from the ACCE framework and evaluating genetic tests in the United Kingdom. *Genetics in Medicine* 7(7), 495-500 (2005).
46. Singer DRJ, Watkins J: Using companion and coupled diagnostics within strategy to personalize targeted medicines. *Personalized Medicine* 9(7), 751-761 (2012).
47. Thariani R, Veenstra DL, Carlson JJ, Garrison LP, Ramsey S: Paying for personalized care: Cancer biomarkers and comparative effectiveness. *Molecular Oncology* 6(2), 260-266 (2012).
48. Van Bebber SL, Keegan HL, Phillips KA, Issa AM: Novel personalized medicine technology: UGT1A1 testing for irinotecan as a case study. *Personalized Medicine* 3(4), 415-419 (2006).
49. Van Den Akker-Van Marle ME, Gurwitz D, Detmar SB et al.: Cost-effectiveness of pharmacogenomics in clinical practice: A case study of thiopurine methyltransferase genotyping in acute lymphoblastic leukemia in Europe. *Pharmacogenomics* 7(5), 783-792 (2006).
50. Van Rooij T, Wilson DM, Marsh S: Personalized medicine policy challenges: Measuring clinical utility at point of care. *Expert Review of Pharmacoeconomics and Outcomes Research* 12(3), 289-295 (2012).
51. Veenstra DL, Higashi MK, Phillips KA: Assessing the cost-effectiveness of pharmacogenomics. *AAPS PharmSciTech* 2(3), E29 (2000).
52. Veenstra DL, Piper M, Haddow JE et al.: Improving the efficiency and relevance of evidence-based recommendations in the era of whole-genome sequencing: An EGAPP methods update. *Genetics in Medicine* 15(1), 14-24 (2013).

Appendix 7: Search strategies used to identify health economic evidence to support current genetic testing practice in CLL

Embase (1988 to present)

chronic lymph\$ leuk?emia.ti,ab.
exp chronic lymphatic leukemia/
or/1-2
cytogenetic\$.ti,ab.
fish\$.ti,ab.
fluorescent in situ.ti,ab.
microarray\$.ti,ab.
dna chip\$.ti,ab.
dna microarray\$.ti,ab.
oligonucleotide microarray\$.ti,ab.
biochip\$.ti,ab.
affymetrix.ti,ab.
illumina.ti,ab.
array cgh.ti,ab.
snp array.ti,ab.
copy number.ti,ab.
17p.ti,ab.
11q.ti,ab.
13q.ti,ab.
trisomy.ti,ab.
genomic complexity.ti,ab.
clonal evolution.ti,ab.
mosaicism.ti,ab.
igvh.ti,ab.
ighv.ti,ab.
immunoglobulin heavy chain.ti,ab.
zap-70.ti,ab.
cd38.ti,ab.
mutat\$.ti,ab.
deletion\$.ti,ab.
genetic\$.ti,ab.
exp genetics/
genomic\$.ti,ab.
exp genomics/
pharmacogenetic\$.ti,ab.
exp pharmacogenetics/
pharmacogenomic\$.ti,ab.
exp pharmacogenomics
or/4-38
economic adj3 evaluation. ti,ab.
exp health economics/ or exp economic evaluation/
cost adj3 effectiveness. ti,ab.

exp "cost effectiveness analysis"/
 cost adj3 benefit. ti,ab.
 exp "cost benefit analysis"/
 cost adj3 utility. ti,ab.
 exp "cost utility analysis"/ or exp "quality of life"/
 cost adj3 minimi#ation. ti,ab.
 exp "cost minimization analysis"
 cost adj3 consequence.ti,ab.
 cost\$. ti,ab.
 exp "health care cost"/ or exp "cost"/
 pharmacoeconomic\$.ti,ab.
 life year\$.ti,ab.
 quality adjusted.ti,ab.
 utility\$.ti,ab.
 benefit.ti,ab.
 or/40-57
 3 and 39 and 58
 limit 59 to english language
 editorial.pt
 letter.pt
 note.pt
 (rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep or primate or primates).ti,ab.
 or/61-64
 60 not 65
 limit 66 to human
 limit 67 to yr="2000 -Current"

Medline In-Process & Other Non-Indexed Citations and Medline 1946 to present + Medline Daily Update

chronic lymph\$ leuk?emia.ti,ab.
 exp Leukemia, Lymphocytic, Chronic, B-Cell/
 or/1-2
 cytogenetic\$.ti,ab.
 fish\$.ti,ab.
 fluorescent in situ.ti,ab.
 microarray\$.ti,ab.
 dna chip\$.ti,ab.
 dna microarray\$.ti,ab.
 oligonucleotide microarray\$.ti,ab.
 biochip\$.ti,ab.
 affymetrix.ti,ab.
 illumina.ti,ab.
 array cgh.ti,ab.
 snp array.ti,ab.
 copy number.ti,ab.
 17p.ti,ab.
 11q.ti,ab.
 13q.ti,ab.
 trisomy.ti,ab.

genomic complexity.ti,ab.
 clonal evolution.ti,ab.
 mosaicism.ti,ab.
 igvh.ti,ab.
 ighv.ti,ab.
 immunoglobulin heavy chain.ti,ab.
 zap-70.ti,ab.
 cd38.ti,ab.
 mutat\$.ti,ab.
 deletion\$.ti,ab.
 genetic\$.ti,ab.
 exp genetics/
 genomic\$.ti,ab.
 exp genomics/
 pharmacogenetic\$.ti,ab.
 exp pharmacogenetics/
 pharmacogenomic\$.ti,ab.
 exp pharmacogenomics
 or/4-38
 economic adj3 evaluation.ti,ab.
 exp Economics, Medical/ or exp Health Care Rationing/
 cost adj3 effectiveness. ti,ab.
 cost adj3 benefit.ti,ab.
 exp cost benefit analysis/
 cost adj3 utility.ti,ab.
 exp "quality of life"/
 cost adj3 minimi#ation.ti,ab.
 cost adj3 consequence.ti,ab.
 cost*. ti,ab.
 exp "costs and cost analysis"/ or exp Health Care Costs/
 pharmacoeconomic*.ti,ab.
 life year*.ti,ab.
 quality adjusted.ti,ab.
 utilit*.ti,ab.
 benefit.ti,ab.
 exp Quality-Adjusted Life Years/
 or/40-56
 3 and 39 and 57
 limit 58 to english language
 editorial.pt
 letter.pt
 (rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or
 cats or bovine or sheep or primate or primates).ti,ab.
 or/60-62
 59 not 63
 limit 64 to human
 limit 65 to yr="2000 -Current"

The Cochrane Library

Searched all text for:

chronic lymphocytic leukaemia AND genetics or genomics

Centre for Reviews and Dissemination

Searched for the following across all 3 databases:

chronic lymphocytic leukaemia [any field] or chronic lymphocytic leukemia [any field]

Appendix 8: Genetic testing for chronic lymphocytic leukaemia in UK NHS laboratories

Laboratory location	Genetic tests offered for chronic lymphocytic leukaemia							
	<i>TP53</i>	<i>ATM</i>	del 13q14	del 13q34	Trisomy 12	14q rearrangement	del 6q	IgV _H
Belfast City Hospital	✓		✓	✓	✓			
Bristol, Southmead Hospital	✓	✓	✓		✓	✓	✓	✓
Cardiff, University Hospital	✓	✓						
Dundee, Ninewells Hospital	✓	✓	✓		✓			
Glasgow Royal Infirmary	✓	✓						
Glasgow, Southern General Hospital	✓	✓	✓	✓	✓			
Leeds, St James's University Hospital	✓	✓	✓					✓
Liverpool Women's NHS Foundation Trust	✓	✓		✓				
London, Great Ormond Street	✓	✓	✓		✓			
London, Hammersmith Hospital	✓	✓	✓	✓	✓			
London, King's College Hospital	✓	✓	✓	✓	✓	✓		
The Royal London Hospital	✓	✓	✓	✓	✓			
Manchester, Christie Hospital	✓	✓	✓ ^a	✓ ^a	✓ ^a		✓ ^a	✓
Nottingham University Hospitals NHS Trust	✓	✓					✓	
Oxford BRC Molecular Diagnostic Laboratory	✓							
Plymouth, Derriford Hospital								✓
Sheffield Children's NHS Foundation Trust	✓	✓		✓				
Sutton, Royal Marsden Hospital	✓	✓	✓	✓	✓			
Total laboratories (%)	17 (94%)	15 (83%)	11 (61%)	9 (50%)	10 (56%)	2 (11%)	3 (17%)	4 (22%)

IgV_H = Immunoglobulin heavy chain variable region. ^a On request only. Sources: Laboratory websites, orhpa.net, personal communications from laboratory directors. Correct to 8th November 2013.

Appendix 9: Survey of haematologists

This appendix reports the methods and results of the survey of haematologists.

Methods

The main aim of this survey was to collect information on genetic and genomic testing in chronic lymphocytic leukaemia (CLL) in the UK National Health Service (NHS). The two specific objectives of this survey were:

- To establish current genetic testing practice in CLL in the UK NHS;
- To identify the opportunities and challenges associated with the translation of new genomic tests into clinical practice in CLL.

Target respondents for the survey were UK haematologists with an interest/specialism in CLL. Potential respondents were informed that if they had ever seen a patient with CLL in an outpatient clinic they were eligible to participate.

Respondents were identified in two ways. First, all members of the National Cancer Research Network CLL subgroup were asked to participate, and to forward the link to colleagues working in haematology and CLL. Second, a link to the survey was circulated to a UK CLL forum mailing list.

A list of questions that reflected the main aims of this survey was developed with the assistance of haematologists based in the Oxford University Hospitals NHS Trust. The final survey was split into five sections. Respondents were initially presented with some introductory text before answering some preliminary questions to establish that they treated CLL patients, and to understand whether they currently had access to a laboratory offering genetic tests for use in routine clinical practice in CLL patients. If respondents did have access to such a laboratory they moved on to Section B, in which they were asked to provide information about the services offered by this laboratory, and their current use of this laboratory. If respondents did not have access to such a laboratory they moved on to Section C, in which they were asked some questions regarding their potential use of such a laboratory, if it existed. All respondents then answered questions in Section D regarding their current and potential future use of genomic testing in CLL. Finally, in Section E information on the personal characteristics of respondents was collected. The final list of questions is provided in **Appendix 10**.

The survey was conducted from January to May 2014 using the LimeSurvey online survey application [194]. Potential respondents were contacted by email to request participation and were given a deadline of one month for completion. Potential respondents received one email reminder. Following the survey end date data was extracted from LimeSurvey and analysed in Microsoft Excel and Stata 12 [240,277]. All responses to open-ended questions were reviewed to identify common themes.

Results

Seventeen haematologists completed the survey. As the number of potential respondents was unknown, a response rate could not be calculated (although it should be noted that there are 34 members of the National Cancer Research Network CLL subgroup).

Table 1 describes respondent characteristics. All see CLL patients in an outpatient clinic and all have access to a laboratory offering genetic testing for CLL patients. Most work in NHS Teaching Hospitals and have been qualified for 15 years. The following sections describe the results with respect to current and future use of genetic and genomic testing.

Table 1: Respondent characteristics

Characteristic	Value	N
Number of respondents who see CLL patients in an outpatient clinic (%)	17 (100%)	17
Number of respondents with access to a laboratory offering genetic testing for CLL patients (%)	17 (100%)	17
Number of respondents who work in a NHS teaching hospital (%)	13 (93%)	14
Mean number of haematologists working in each NHS Trust (SD)	12.6 (7.5)	14
Mean number of haematologists working in each Trust who see CLL patients (SD)	5.6 (5.4)	14
Mean length of time since qualification as a haematologist, in years (SD)	14.6 (9.5)	14

CLL = chronic lymphocytic leukaemia; NHS = national health service; SD = standard deviation.

Current access to genetic testing for chronic lymphocytic leukaemia

All respondents were familiar with the range of genetic tests offered to CLL patients in their nearest laboratory. All laboratories offer fluorescent in situ hybridisation (FISH) testing, nine (53%) offer immunoglobulin heavy chain variable region (IgV_H) mutation analysis, six (35%) offer Sanger sequencing and one (6%) offers karyotyping. **Table 2** provides information on the abnormalities that can be identified by each test.

Table 2: Availability of genetic testing, by abnormality and genetic test

Abnormality	Number of respondents (n=17) who reported that genetic tests for specific abnormalities were available, by genetic test (%)		
	FISH testing	Sanger sequencing	Karyotyping
17p deletions/TP53 losses	14 (82%)	5 (29%)	0
11q deletions/ATM losses	14 (82%)	1 (6%)	0
13q14 deletions	13 (76%)	0	0
13q34 deletions	6 (35%)	0	0
Trisomy 12	13 (76%)	0	1 (6%)
14q rearrangements	6 (35%)	0	1 (6%)
6q deletions	5 (29%)	0	1 (6%)
Other	0	1 (6%) ^a	0
Don't know	0	1 (6%)	0

FISH = fluorescent in situ hybridisation. ^a SF3B1 mutations.

Most respondents reported good access to FISH testing for the main abnormalities, although no respondent reported access to FISH testing for all key abnormalities (17p deletions, 11q deletions, 13q14 deletions, trisomy 12). Access to Sanger sequencing and karyotyping was more patchy.

Current use of genetic testing for chronic lymphocytic leukaemia

15 respondents (88%) reported that they request genetic tests as part of the clinical evaluation of CLL patients at diagnosis. The mean proportion of patients in which these tests are requested at diagnosis is 58% (SD: 39%). All respondents reported that they request genetic tests as part of the clinical evaluation of CLL patients prior to starting treatment. The mean proportion of patients in which these tests are requested prior to starting treatment is 84% (SD: 28%). **Table 3** describes the proportion of respondents who currently use each type of test at diagnosis or prior to starting treatment. FISH testing is used by almost all respondents, while IgV_H mutation analysis is used by around a third of respondents at diagnosis, but infrequently prior to starting treatment. Sanger sequencing is used infrequently at diagnosis but by around a third of respondents prior to starting treatment.

Table 3: Current genetic testing practice at diagnosis and prior to starting treatment

	Number of respondents (%) using each type of test (n=17)			
	FISH testing	IgV _H mutation analysis	Sanger sequencing	Other
Requested at diagnosis	15 (88%)	6 (35%)	2 (12%)	0
Requested prior to starting treatment	16 (94%)	2 (12%)	5 (29%)	1 ^a

FISH = fluorescent in situ hybridisation; IgV_H = immunoglobulin heavy chain variable region. ^a Test not specified.

Table 4 provides further detail on the findings in **Table 3**, describing the abnormalities that are commonly tested at diagnosis and prior to starting treatment, by type of genetic test.

Table 4: Current genetic testing practice at diagnosis, by abnormality

Abnormality	Number of respondents (%) requesting information on each abnormality, by test (n=17)			
	FISH testing		Sanger sequencing	
	At diagnosis	Prior to starting treatment	At diagnosis	Prior to starting treatment
17p deletions/TP53 losses	12 (71%)	13 (76%)	1 (6%)	4 (24%)
11q deletions/ATM losses	10 (59%)	11 (65%)	0	1 (6%)
13q14 deletions	9 (53%)	9 (53%)	1 (6%)	2 (12%)
13q34 deletions	4 (24%)	4 (24%)	0	1 (6%)
Trisomy 12	9 (53%)	9 (53%)	1 (6%)	2 (12%)
14q rearrangements	6 (35%)	4 (24%)	1 (6%)	0
6q deletions	5 (29%)	3 (18%)	0	0
Other	1 (6%) ^a	0	1 (6%) ^b	0
Don't know	0	0	0	0

FISH = fluorescent in situ hybridisation. ^a SF3B1 mutations; ^b SF3B1 and NOTCH1 mutations.

In terms of FISH testing, over half of all respondents request information on each of the four key abnormalities (deletions of 17p, 11q, 13q14 and Trisomy 12) at diagnosis. Information on 6q and 13q34 deletions, and 14q rearrangements is requested by around a third of respondents. Prior to starting treatment, information on 17p and 11q deletions is requested slightly more frequently, while information on 14q rearrangements and 6q deletions is requested slightly less frequently. Information from Sanger sequencing is requested infrequently at diagnosis but slightly more often prior to starting treatment (predominantly following a FISH test which indicates a 17p deletion).

Respondents were asked to describe the clinical utility of possessing this genetic information at diagnosis and prior to starting treatment. **Table 5** summarises these responses. Eight responses (53%) noted that possessing this information at diagnosis can inform discussion of prognosis. Other respondents commented that possessing this information at diagnosis is useful to guide future treatment decisions, even if it isn't acted on straight away. Almost all respondents stated that the main benefit of possessing genetic information prior to starting treatment is to allow patients to be stratified into different treatment groups according to their likely response to therapy. 17p deletions (or *TP53* losses) were the most commonly cited abnormality.

Table 5: What is the clinical utility of possessing genetic information?

Respondent ID	...at diagnosis?	...prior to starting treatment?
s11	Assessment of risk of progression, transformation, refractoriness	Refractoriness to chemo, risk of transformation
s12	IgV _H mutation analysis helpful in predicting likely progression	Information on p53 deletion/mutation vital for choice of therapy
s13	Prognostic information	-
s14	Prognostic significance	Need information on p53 deletion/mutation to stratify treatment
s15	TP53 deletion (ideally mutation as well should be detected which we don't test for) will make us change our treatment i.e. high risk trial or campath/methylprednisolone	As before only p53 deletion is useful to know prior to treatment
s17	Informs discussion of prognosis and in case of TP53 deletion, form of treatment when required	To assess whether p53 deletion present to assess whether to treat with standard chemotherapy +/- monoclonal antibody or to use steroids etc. if p53 deletion.
s18	To direct therapy	Direct therapy
s19	None. Mainly for research purposes	Selection of TP53 independent therapy in patients with deletions/mutations TP53
s20	-	See guidelines
s23	Only if immediate treatment needed. Very occasionally more tests performed in patients who wish to have full prognostic work up although this is not routine	Guides treatment
s25	-	p53 deletion influences treatment choice
s26	I don't perform at diagnosis. I perform at time of need for treatment	Selection of treatment, likely outcomes
s27	Prognostic info	Choice of treatment
s32	Guide to prognosis	Guides therapy and predicts chemo-refractoriness
s33	To decide on treatment suitability e.g. standard chemo vs MabCampath	To consider chemotherapy: standard chemo vs MabCampath
s39	Some patients at diagnosis require immediate treatment. TP53 mutations low incidence but will test in young patient with high counts etc.	Occasionally also get additional tests back from lab. Main reason for testing is seek P53 deletion/mutation
s43	If likely to require treatment at presentation sensitivity to purine analogues is tested	Chemotherapy resistance, prognosis

Future use of genomic testing in chronic lymphocytic leukaemia

Ten of the 16 respondents (63%) who answered questions in this section reported that they had requested a genomic test for a CLL patient for research purposes. Six (38%) had requested targeted next generation sequencing, 3 (19%) had requested whole genome sequencing, and 5 (31%) had requested whole exome sequencing. Respondents were asked to describe the clinical utility of possessing the information provided by this test and to indicate how the test results impacted on the management of CLL in this patient. Responses are listed in **Table 6**. The general feeling was that these techniques did not yet provide sufficient clinical utility to be translated into current practice.

Table 6: What is the clinical utility of possessing genomic testing information?

Respondent ID	Response
s11	Still too many false negative by NGS techniques
s12	Little (utility) at the moment
s13	Difficult to predict because of changing treatment paradigms
s15	Research questions relating to validation of target
s19	Research only - no clinically useful benefit at present
s20	None. Research value only currently
s32	Treatment choice
s33	Did not impact
s43	P53 mutation/deletion less likely to respond to FCR and relapse quickly - more resistant disease and require other treatment

Respondents were then presented with descriptions of two hypothetical genomic tests for CLL patients which could be requested either at diagnosis or prior to beginning treatment. Respondents were asked to make the simplifying assumption that the genetic tests that are currently available (usually FISH testing and/or Sanger sequencing) can identify around one-third of those CLL patients who will not respond to usual chemotherapy treatment. For each hypothetical test, respondents were asked to indicate the percentage of all CLL patients for whom they would request this test at diagnosis and prior to starting treatment (**Tables 7-10**). Respondents were also asked to indicate to what extent factors such as test cost, time taken to receive test results, test sensitivity or test specificity influenced their decision to request each test (**Table 11**).

Test One: Assume that a genomic test is available that could identify twice as many (i.e. two-thirds) of those CLL patients who will not respond to usual chemotherapy treatment. Please indicate the percentage of all CLL patients for whom you would request this test at diagnosis.

Respondents gave a mixed response to Test One. Some noted the additional prognostic information provided by such a test, but many preferred to only use this test prior to starting treatment. This test would only be used in 50% of patients overall.

Table 7: Percentage of chronic lymphocytic leukaemia patients for whom respondents would recommend Test One, and why

Respondent ID	%	Explanation
s11	100	CLL is very heterogeneous
s12	100	That would be hugely helpful in terms of FCR, allowing us to switch to more targeted therapies where traditional prognostic markers no longer apply
s14	10	Only indicated in those in whom I might be considering treatment soon
s15	30	Would only test patient when I knew that treatment was likely to be required
s17	40	Unless additional testing has independent prognostic value, I would consider these tests prior to treatment rather than at diagnosis
s18	100	The question is too simplistic given the rapidly changing therapeutic arena in CLL. Obviously knowing if patients have a p53 abnormality is very important when commencing therapy today but in the future when signalling pathway inhibitors are the mainstay of therapy, knowing the p53 status will diminish in importance. You may only need to know a patients' p53 status when signalling inhibitor therapy fails
s19	0	Not useful to have a marker predictive of treatment response if patient does not require treatment
s20	100	Presumably will identify patients less likely to do as well which may affect management decisions - particularly if toxicity issues with treatment
s23	0	Would not inform management
s25	0	Doesn't change initial management. would do pre-treatment
s26	0	Would only test pre-treatment
s27	95	Treatment choice
s32	100	This informs future treatment options and prognosis
s33	50	The test would be useful for patients able to tolerate alternative treatments
s39	60	Depends on cost and efficacy. Not likely to test old patients
s43	10	Current prescribing restrictions mean that we would need to treat and reassess before changing treatment. In a clinical trial we would expect additional molecular information but not have the results
Mean % (SD)	50 (43)	

Test Two: Assume that a genomic test is available that could identify twice as many (i.e. two-thirds) of those CLL patients who will not respond to usual chemotherapy treatment. Please indicate the percentage of all CLL patients for whom you would request this test prior to beginning treatment.

Test Two was more popular with respondents. Two-thirds of respondents would use this test in all patients, and it would be used in 85% of patients overall. Respondents overwhelmingly cited improved treatment stratification as justification for their responses.

Table 8: Percentage of chronic lymphocytic leukaemia patients for whom respondents would recommend Test Two, and why

Respondent ID	%	Explanation
s11	100	Heterogeneity of CLL and implications for prognosis
s12	100	Why not do it? Cost, I suppose, but savings in terms of toxicities of ineffective therapy would easily outweigh additional costs?
s14	100	Would use this to stratify treatment
s15	80	Would be extremely useful to avoid ineffective and toxic therapies
s17	95	If these tests were available I would perform them in all patients except those who may be unfit for any alternate treatment
s18	100	Please refer to my previous answer
s19	100	May help to inform selection of most appropriate therapy
s20	100	As before
s23	20	No info as to how this would guide therapy
s25	100	Influences treatment choice
s26	100	Helpful in selecting treatment
s27	100	Treatment choice
s32	100	Impact treatment options
s39	60	See previous
s43	20	Majority are older, frail and not candidates for more immunosuppressive chemotherapy
Mean % (SD)	85 (29)	

Test Three: Assume that a genomic test is available that could identify twice as many (i.e. two-thirds) of those CLL patients who will not respond to usual chemotherapy treatment, and provide information on genomic complexity and clonal evolution. Please indicate the percentage of all CLL patients for whom you would request this test at diagnosis.

Test Three was the least favoured test, and would be used in only a third of patients. Respondents disputed the clinical relevance of this additional information, doubting that it would change patient management. However, it was noted that this information may influence the frequency of monitoring.

Table 9: Percentage of chronic lymphocytic leukaemia patients for whom respondents would recommend Test Three, and why

Respondent ID	%	Explanation
s11	0	No clear dominant role of genomic complexity and clonal evolution on management of patients and on treatment choice
s12	0	Uncertain about clinical relevance of the other two parameters
s14	10	No different from previous question - genetic stability a research question
s15	30	Same as before its really most useful in patients requiring therapy. I doubt it would be useful as a predictor for stage A patients?
s17	60	This information may influence frequency of monitoring
s18	20	Around 20% of my newly diagnosed patients go on to receive therapy immediately
s19	50	Only in those patients where this information is likely to be helpful in future management
s20	100	I would like the data
s25	0	Doesn't currently influence management
s26	0	Would only test at time of treatment
s27	25	Clinical utility unknown
s32	100	Inform prognosis
s39	60	Once I embark on treatment I like to know what kind of CLL they have
s43	10	Those patients who require treatment at diagnosis
Mean % (SD)	33 (35)	

Test Four: Assume that a genomic test is available that could identify twice as many (i.e. two-thirds) of those CLL patients who will not respond to usual chemotherapy treatment, and provide information on genomic complexity and clonal evolution. Please indicate the percentage of all CLL patients for whom you would request this test prior to beginning treatment.

Test Four would be used in 58% of patients. Again, respondents were unsure if it would change clinical practice, but would be more willing to use it further along clinical pathways prior to treatment (as compared to Test Three).

Table 10: Percentage of chronic lymphocytic leukaemia patients for whom respondents would recommend Test Four, and why

Respondent ID	%	Explanation
s11	0	No real demonstration that genomic complexity is an independent prognostic factor
s12	0	Well I guess it depends on what you mean by clonal evolution - if you could detect low levels of p53 mutant cells that would be helpful downstream - but this would not alter therapy
s14	100	Very valuable information and could save exposure to agents to which patients would not respond
s15	80	As previous answers
s17	95	Reasoning as before - I would perform this test in all patients in whom it would change management.
s18	0	The question is really a research question not a clinically relevant NHS question. What different therapy would you give if you knew a patient was likely to get clonal evolution - no robust data to suggest any therapeutic approach is any better than any other
s19	75	In patients where the additional information may be important in planning treatment
s20	100	I think the same answers apply broadly to all these questions. Although I would really like this information, in all seriousness, we will have to show how it is useful to direct patient care. For all of these questions, I think about IgV _H sequencing - we have access to this data, but how many clinicians actually use it in routine practice? However, if more detailed genomic analysis will help us separate out which patients will benefit most from which therapy, then the incentive to use these tests will increase significantly
s25	50	If this doesn't influence clinical management then it is probably best done as a research tool
s26	100	Helpful for selecting treatment
s27	50	Clinical utility unknown
s32	100	Impact treatment and prognosis
s39	60	Not needed in very elderly
s43	0	Prognostic factors may change with disease evolution
Mean % (SD)	58 (42)	

Overall, respondents showed a stronger preference for genomic testing prior to treatment, rather than at diagnosis, and were more convinced of the clinical relevance of identifying patients who would not respond to chemotherapy than of identifying patients with tumours exhibiting genomic complexity or clonal evolution.

Various factors influenced the decision to request each of these genomic tests. In general, test sensitivity and specificity were more important than test cost or time taken to receive test result,

suggesting that improvements in these test parameters would have the greatest impact on the use of genomic testing in the future.

Table 11: Influence of various factors on the decision to request a genomic test

Test	N	Influence, by factor (1 = no influence; 10 = highly influential): Mean (SD)				
		Test cost	Time to test result	Sensitivity	Specificity	Other
Test One	16	8.5 (2.1)	5.4 (3.3)	8.1 (1.1)	8.4 (1.0)	8.6 (1.3) ^a
Test Two	14	7.8 (2.5)	6.9 (3.0)	8.5 (1.2)	8.8 (1.1)	8.7 (1.2) ^b
Test Three	14	7.9 (2.9)	5.9 (3.0)	8.2 (1.2)	8.6 (1.1)	10.0 (N/A) ^c
Test Four	14	7.5 (3.1)	6.4 (3.1)	7.6 (2.5)	8.0 (2.5)	-

^a Seven respondents identified other factors, which were 'patient consent', 'good data supporting changing treatment based on result', 'proven clinical utility/benefit', 'if it was strongly predictive of time to progression', 'no impact on treatment, hard to use information', 'only 50% of patients with low counts progress', and 'influence of treatment sensitivity'; ^b Three respondents identified other factors, which were 'patient choice', 'MDT', and 'the data supporting such a strategy to alter treatment based on this result'; ^c One patient identified another factor, which was 'patient consent'.

Appendix 10: List of questions in the haematologist survey

Introductory text

Welcome to our survey on genetic and genomic testing in chronic lymphocytic leukaemia (CLL) in the UK NHS.

This survey has two aims:

- To establish current genetic testing practice in CLL.
- To identify the opportunities and challenges associated with the translation of new genomic tests into clinical practice in CLL.

This survey is being run amongst all UK haematologists with an interest/specialism in CLL. **If you have ever seen a patient with CLL in an outpatient clinic, you are eligible to take part in our survey.**

We expect that this survey will take 7-8 minutes to complete, on average.

Questions marked with a red asterisk (*) are compulsory. All responses will remain 100% anonymous.

It is important that as many UK haematologists as possible reply, as your responses will inform the translation of new genetic and genomic tests for CLL from research into clinical practice in the UK NHS.

This survey is part of a larger project of work investigating the methodological challenges associated with conducting health economic evaluations of genomic tests in the UK NHS. This work is being undertaken at the Health Economics Research Centre (HERC) at the University of Oxford. More information can be found at <http://bit.ly/ZJqSkI>.

THANK YOU IN ADVANCE FOR YOUR TIME AND COOPERATION

Questions

Section A – Preliminary questions

1. Do you ever see patients with chronic lymphocytic leukaemia (CLL) in an outpatient clinic?
 - Yes (go to Q2)
 - No (go to section E)

Genetic tests currently exist which can provide information to inform disease management in CLL patients. These tests can take various forms e.g. fluorescence in situ hybridisation (FISH) testing or Sanger sequencing. These tests can provide a variety of information, such as the presence of *TP53* mutations.

2. Do you currently have access to a laboratory offering genetic tests for use in routine clinical practice in CLL patients?
 - Yes (go to Section B)
 - No (go to Section C)

END OF SECTION A

Section B: Current practice - access to a laboratory offering genetic tests for CLL

3. Please give the name and location of this laboratory.
 - Free text box
4. Do you know which genetic tests are offered by this laboratory for use in routine clinical practice in CLL patients?
 - Yes (go to Q5)
 - No (go to Q7)
5. Please indicate which genetic tests are offered by this laboratory for use in routine clinical practice in CLL patients. If a test is offered that is not in this list, please provide details in the 'Other' box.
 - Fluorescence in situ hybridisation (FISH) testing
 - Immunoglobulin variable region heavy chain (IgVH) mutation analysis
 - Sanger sequencing
 - Other [free text box]
6. Please indicate what genetic information is provided by fluorescence in situ hybridisation (FISH) testing in this context. [Question repeated for Sanger sequencing and Other tests]
 - TP53 mutations
 - ATM mutations
 - del 13q14
 - del13q34
 - Trisomy 12
 - 14q rearrangement
 - del 6q
 - I don't know
 - Other [free text box]
7. Do you ever request genetic tests as part of the clinical evaluation of CLL patients at diagnosis?
 - Yes (go to Q8)
 - No (go to Q12)
8. In what percentage of CLL patients do you request these tests at diagnosis?
 - Enter percentage (plus 'unknown' option)
9. Please indicate which genetic tests are requested at diagnosis. If 'Other' is selected, it is assumed that this is the same test specified earlier in the survey..
 - Same options as Q5
10. Please indicate what genetic information is provided by fluorescence in situ hybridisation (FISH) testing in this context at diagnosis. [Question repeated for Sanger sequencing and Other tests]
 - Same options as Q6
11. Please describe the clinical utility of possessing this information at diagnosis.
 - Free text box (then go to Q13)

12. Why are genetic tests not requested as part of the clinical evaluation of CLL patients at diagnosis?
- Free text box
13. Do you ever request genetic tests prior to beginning treatment in CLL patients?
- Yes (go to Q14)
 - No (go to Q18)
14. In what percentage of CLL patients do you request these tests prior to beginning treatment?
- Enter percentage (plus 'unknown' option)
15. Please indicate which genetic tests are requested prior to beginning treatment. If 'Other' is selected, it is assumed that this is the same test specified earlier in the survey..
- Same options as Q5
16. Please indicate what genetic information is provided by fluorescence in situ hybridisation (FISH) testing in this context prior to beginning treatment. [Question repeated for Sanger sequencing and Other tests]
- Same options as Q6
17. Please describe the clinical utility of possessing this information prior to beginning treatment.
- Free text box (then go to Section C)
18. Why are genetic tests not requested prior to beginning treatment in CLL patients?
- Free text box (then go to Section C)

END OF SECTION B

Section C: Current practice - no access to a laboratory offering genetic tests for CLL

19. If you had access to a laboratory offering these tests, would you request them as part of the clinical evaluation of CLL patients at diagnosis?
- Yes (go to Q20)
 - No (go to Q22)
20. What genetic information would you request at diagnosis, if you had access to a laboratory offering these tests for CLL patients?
- Same options as Q5
21. Please describe the clinical utility of possessing this information at diagnosis.
- Free text box
22. If you had access to a laboratory offering these tests, would you request them prior to beginning treatment in CLL patients?
- Yes (go to Q23)
 - No (go to Section D)
23. What genetic information would you request prior to beginning treatment, if you had access to a laboratory offering genetic tests for CLL patients?
- Same options as Q5

24. Please describe the clinical utility of possessing this information prior to beginning treatment.
- Free text box (then go to Section D)

END OF SECTION C

Section D: Future practice

Genomic tests are currently being developed that can provide additional information to inform disease management in CLL patients, over and above that provided by genetic tests. For example, these tests may be able to identify more of the patients with *TP53* mutations, or track clonal evolution and genomic complexity. These tests can take various forms, for example:

- Microarray analysis;
- Targeted next generation sequencing;
- Whole genome sequencing;
- Whole exome sequencing.

At present, these tests are usually only available for research purposes. However, it is likely that they will become available for routine clinical use within the next 5 years.

25. Have you ever requested a genomic test for a CLL patient for research purposes?

- Yes (go to Q26)
- No (go to Q29)

26. Do you know what type of test you requested?

- Yes (go to Q27)
- No (go to Q29)

27. Please indicate which test(s) you requested.

- Microarray analysis
- Targeted next generation sequencing
- Whole genome sequencing
- Whole exome sequencing
- Other genomic test: please describe (free text box)
- Not known

28. Please describe the clinical utility of possessing the information provided by this test and indicate how the test results impacted on the management of CLL in this patient (or these patients).

- Free text box

We would now like you to answer some questions about two hypothetical genomic tests, which could be requested **at diagnosis** or **prior to beginning treatment**. For the purposes of these questions, we would like you to make the simplifying assumption that the genetic tests that are currently available (usually FISH testing and/or Sanger sequencing) can identify around one-third of those CLL patients who will not respond to usual chemotherapy treatment.

29. Assume that a genomic test is available that could identify twice as many of these patients (i.e. two-thirds of those CLL patients who will not respond to usual chemotherapy treatment). Please indicate the percentage of all CLL patients for whom you would request this test at diagnosis?
- Enter percentage
30. Please explain your response.
- Free text box
31. Please indicate to what extent each of the following factors would influence your decision to request this genomic test for a CLL patient at diagnosis (ranking scale to click on for each factor: 0 = no influence; 10 = highly influential).
- Test cost
 - Time taken to receive test result
 - Sensitivity of the test
 - Specificity of the test
 - Other factors (please state – free text box)
32. Assume that a genomic test is available that could identify twice as many of these patients (i.e. two-thirds of those CLL patients who will not respond to usual chemotherapy treatment). Please indicate the percentage of all CLL patients for whom you would request this test prior to beginning treatment?
- Enter percentage
33. Please explain your response.
- Free text box
34. Please indicate to what extent each of the following factors would influence your decision to request this genomic test for a CLL patient prior to beginning treatment.
- Same options as Q31
35. Assume that a genomic test is available that could identify twice as many of these patients (i.e. two-thirds of those CLL patients who will not respond to usual chemotherapy treatment) and provide information on genomic complexity and clonal evolution. Please indicate the percentage of all CLL patients for whom you would request this test at diagnosis?
- Enter percentage
36. Please explain your response.
- Free text box
37. Please indicate to what extent each of the following factors would influence your decision to request this genomic test for a CLL patient at diagnosis.
- Same options as Q31
38. Assume that a genomic test is available that could identify twice as many of these patients (i.e. two-thirds of those CLL patients who will not respond to usual chemotherapy treatment) and provide information on genomic complexity and clonal evolution. Please indicate the percentage of all CLL patients for whom you would request this test prior to beginning treatment?
- Enter percentage

39. Please explain your response.

- Free text box

40. Please indicate to what extent each of the following factors would influence your decision to request this genomic test for a CLL patient **prior to beginning treatment**.

- Same options as Q31 (then go to Section E)

END OF SECTION D

Section E: Personal information

We would now like to ask some questions about you. Your responses will help us to understand whether different subgroups of haematologists have different opinions about genetic and genomic testing. All of your answers will be kept confidential.

41. Do you work in an NHS teaching hospital?

- Yes/No

42. Including yourself, how many haematologists are employed within your Trust?

- Enter number (plus 'unknown' option)

43. Including yourself (if applicable), how many haematologists see patients with CLL within your Trust?

- Enter number (plus 'unknown' option)

44. For how many years have you been working as a qualified consultant haematologist?

- Enter number of years

45. Are there any comments you would like to make regarding this survey?

- Free text box

END OF SURVEY

End message

You have reached the end of the survey. Thank you for taking the time to participate. If you would like to be notified about the survey results, please email James Buchanan (james.buchanan@dph.ox.ac.uk) who will add you to our mailing list. You will only receive communications related to this study.

More information about this work can be found on the Health Economics Research Centre website at <http://bit.ly/ZJqSkI>

Appendix 11: Search strategies used to identify health economic evidence for the use of targeted next generation sequencing in all disease areas

Embase (1988 to present)

economic evaluation. ti,ab.
exp economic evaluation/
cost effectiveness. ti,ab.
exp "cost effectiveness analysis"/
cost benefit. ti,ab.
exp "cost benefit analysis"/
cost utility. ti,ab.
exp "cost utility analysis"/
cost minimi#ation. ti,ab.
exp "cost minimization analysis"
cost consequence.ti,ab.
or/1-11
next generation sequencing.ti,ab.
whole exome sequencing.ti,ab.
whole genome sequencing.ti,ab.
genome wide technologies.ti,ab.
massively parallel sequencing.ti,ab.
next generation technology.ti,ab.
high throughput sequencing.ti,ab.
pyrosequencing.ti,ab.
ion torrent.ti,ab.
miseq.ti,ab.
illumina sequencing.ti,ab.
solexa sequencing.ti,ab.
or/13-24
12 and 25
limit 26 to english language
editorial.pt
letter.pt
note.pt
(rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep or primate or primates).ti,ab.
or/28-31
27 not 32
limit 33 to human
limit 34 to yr="2005 -Current"

Medline In-Process & Other Non-Indexed Citations and Medline 1946 to present + Medline Daily Update

economic evaluation. ti,ab.
exp Economics, Medical/ or exp Health Care Rationing/

cost effectiveness. ti,ab.
 cost benefit. ti,ab.
 exp cost benefit analysis/
 cost utility. ti,ab.
 cost minimi#ation. ti,ab.
 cost consequence.ti,ab.
 or/1-8
 next generation sequencing.ti,ab.
 whole exome sequencing.ti,ab.
 whole genome sequencing.ti,ab.
 genome wide technologies.ti,ab.
 massively parallel sequencing.ti,ab.
 next generation technology.ti,ab.
 high throughput sequencing.ti,ab.
 pyrosequencing.ti,ab.
 ion torrent.ti,ab.
 miseq.ti,ab.
 illumina sequencing.ti,ab.
 solexa sequencing.ti,ab.
 or/10-21
 9 and 22
 limit 23 to english language
 editorial.pt
 letter.pt
 (rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or
 cats or bovine or sheep or primate or primates).ti,ab.
 or/25-27
 24 not 28
 limit 29 to human
 limit 30 to yr="2005 -Current"

Econlit

(ti,ab(economic evaluation) OR ti,ab,pub(economic*) OR ti,ab(cost) OR ti,ab(cost effective*) OR
 ti,ab(cost utility) OR ti,ab(cost consequence) OR ti,ab(cost minimi?ation) OR ti,ab(cost benefit))
 AND (ti,ab(next generation sequencing) OR ti,ab(whole exome sequencing) OR ti,ab(whole genome
 sequencing) OR ti,ab(genome wide technologies) OR ti,ab(massively parallel sequencing) OR
 ti,ab(next generation technology) OR ti,ab(high throughput sequencing) OR ti,ab(pyrosequencing)
 OR ti,ab(ion torrent) OR ti,ab(miseq) OR ti,ab(illumina sequencing) OR ti,ab(solexa sequencing))

Search limits: English language; limited to 2005 onwards

The Cochrane Library

Searched for:

economic evaluation or cost benefit or cost utility or cost effectiveness or cost minimisation or cost
 consequence [all text]

AND

“next” generation sequencing or whole exome sequencing or whole genome sequencing or genome wide technologies or massively parallel sequencing or “next” generation technology or high throughput sequencing or pyrosequencing or ion torrent or miseq or illumina sequencing or solexa sequencing [title, abstract, keywords]

Centre for Reviews and Dissemination

Searched for the following across all 3 databases:

economic evaluation or cost benefit or cost utility or cost effectiveness or cost minimisation or cost consequence AND next generation sequencing or whole exome sequencing or whole genome sequencing or genome wide technologies or massively parallel sequencing or next generation technology or high throughput sequencing or pyrosequencing or ion torrent or miseq or illumina sequencing or solexa sequencing

Appendix 12: Summary of previous economic evaluations in chronic lymphocytic leukaemia conducted by the National Institute for Health and Care Excellence

Technology appraisal number (year)	Submission title	Modelling approach	Comments on modelling approach from ERG	Sources
TA119 (2007)	Fludarabine phosphate for the first-line treatment of chronic lymphocytic leukaemia	• Markov model • 260 x 28-day cycles (effectively lifetime) • Modelled patients from start of first-line treatment to death • Patient cohort included all untreated patients	• Modelling approach is appropriate • Some concern that variations in second-line treatment in different clinical settings were not fully considered	[221,313,314]
TA174 (2009)	Rituximab for the first-line treatment of chronic lymphocytic leukaemia	• Markov model • Three disease states • Cycle length: 1 month • Time horizon: 15 years • Modelled patients from start of first-line treatment to death • Patient cohort included all untreated patients	• Modelling approach is appropriate	[222,315,316]
TA193 (2010)	Rituximab for the treatment of relapsed/refractory chronic lymphocytic leukaemia	• Markov model • Same structure as TA174 • Time horizon: 25 years • Patients modelled from initiation of relapsed/refractory treatment to death • Patient cohort included all patients requiring relapsed/refractory treatment	• Modelling approach is appropriate	[223,237,317]

Technology appraisal number (year)	Submission title	Modelling approach	Comments on modelling approach from ERG	Sources
TA202 (2010)	Ofatumumab (Arzerra®) for the treatment of chronic lymphocytic leukaemia in patients who are refractory to fludarabine and alemtuzumab	• Partitioned-survival model (similar to a Markov model) • Three disease states • Cycle length: 1 day • Time horizon: 10 years • Patients modelled from initiation of relapsed/refractory treatment to death • Patient cohort included all patients requiring relapsed/refractory treatment	• Modelling approach is appropriate	[197,318,319]
TA216 (2011)	Bendamustine for the first-line treatment of chronic lymphocytic leukaemia	• Markov model • Complex structure: 39 health states • Cycle length: 3 months • Time horizon: 35 years • Modelled patients from start of first-line treatment to death • Patient cohort included only those patients who are considered unfit for fludarabine-based therapies	• Modelling approach is generally appropriate, although there was some concern that the degree of sophistication was not justified by the quality of the data	[224,238,320]

ERG = evidence review group.

Appendix 13: Descriptions of all model health states

State	Description
FCR treatment (first-line, first cycle)	Patients undergo a 28-day cycle of FCR treatment, guided by genetic or genomic testing in some cases. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
FCR treatment (second-line, first cycle)	Patients undergo a 28-day cycle of FCR treatment, guided by genetic testing in some cases. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
FCR treatment (first- or second-line, subsequent cycles)	Patients undergo a further 28-day cycle of FCR treatment. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
BR treatment (first-line, first cycle)	Patients undergo a 28-day cycle of BR treatment, guided by genetic testing in some cases. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
BR treatment (second-line, first cycle)	Patients undergo a 28-day cycle of BR treatment, guided by genetic or genomic testing in some cases. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
BR treatment (first- or second-line, subsequent cycles)	Patients undergo a further 28-day cycle of BR treatment. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
Ofa treatment (first-line, first cycle)	Patients undergo a 28-day cycle of Ofa treatment, guided by genetic or genomic testing. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
Ofa treatment (first-line, second cycle)	Patients undergo a further 28-day cycle of Ofa treatment. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
Ofa treatment (first-line, subsequent cycles)	Patients undergo a further 28-day cycle of Ofa treatment. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
Ofa treatment (second-line or refractory, first cycle)	Patients undergo a 28-day cycle of Ofa treatment, guided by genetic testing in some cases. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
Ofa treatment (second-line or refractory, second cycle)	Patients undergo a further 28-day cycle of Ofa treatment. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
Ofa treatment (second-line or refractory, subsequent cycles)	Patients undergo a further 28-day cycle of Ofa treatment. All receive concomitant medications and periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
Ibrutinib treatment (first-line)	Patients undergo a 28-day cycle of ibrutinib treatment, guided by genomic testing in some cases. All periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.
Ibrutinib treatment (second-line or refractory)	Patients undergo a 28-day cycle of ibrutinib treatment, guided by genetic or genomic testing. All periodically undergo diagnostic tests and attend outpatient appointments. Some patients require blood transfusions.

State	Description
Remission following FCR treatment (all lines)	Patients are in remission for a 28-day cycle. Some take concomitant medications or undergo IV immunoglobulin treatment. All have periodic outpatient appointments and FBCs.
Remission following BR treatment (all lines)	Patients are in remission for a 28-day cycle. Some undergo IV immunoglobulin treatment. All have periodic outpatient appointments and FBCs.
Remission following Ofa treatment (all lines)	Patients are in remission for a 28-day cycle. All have periodic outpatient appointments and FBCs.
Undergoing BMT	Patient has a BMT. This is assumed to take up one 28-day cycle.
Remission following BMT	Patients are in remission following BMT for a 28-day cycle. All have periodic consultant and nurse outpatient appointments, and some telephone consultations. Various tests are performed periodically, and some patients require IV immunoglobulin treatment, donor lymphocyte infusions or blood and platelet support.
BSC	Patients spend a 28-day cycle undergoing treatment which is aimed at managing symptoms. All will have an outpatient appointment and FBC. Some will require blood transfusions or an inpatient stay.

BMT = bone marrow transplant; BR = bendamustine and rituximab; BSC = best supportive care; FBC = full blood count; FCR = fludarabine, cyclophosphamide and rituximab; IV = intravenous; Ofa = ofatumumab.

Appendix 14: Transition probabilities used in the Markov model

P	Disease state	Transition	TP	Notes	Source
p1	Undergoing first-line FCR treatment (no <i>TP53</i> mutation)	Dead from disease	0.0027	The mortality rate for this patient subgroup was reported to be 0.10 (39/386) at three years. This rate was converted into a 28-day transition probability (TP).	[230]
p2	Undergoing first-line FCR treatment (no <i>TP53</i> mutation)	No response to treatment – enters second-line treatment	0.0016	The non-response rate for this patient subgroup was reported to be 0.08 (32/386) at three years. This was converted to a 28-day TP. It was then assumed that a proportion of those patients who do not respond to treatment are those in whom a high risk genetic mutation has emerged. This TP (p3) was therefore subtracted from the TP calculated from Hallek <i>et al.</i> [230].	[230]
p3	Undergoing first-line FCR treatment (no <i>TP53</i> mutation)	High risk genetic mutation emerges – enters refractory therapy	0.0006	Both references reported the proportion of patients who acquired a cytogenetic abnormality that changed their risk classification to a higher level. In Shanafelt <i>et al.</i> [321] this was 5/108 patients (median time from baseline to second sample 5.6 years). In Gunnarsson <i>et al.</i> [322] this was 3/59 patients (median time 6.67 years). This information was used to calculate a 28-day TP for each study (0.0006 in both cases). An unweighted average TP was then calculated. In the absence of any information to the contrary in the literature, the same TP was assumed to apply to patients receiving either FCR or BR treatment, and also patients identified by genomic testing (correctly or otherwise) to be likely FCR responders.	[321,322]
p4	Undergoing first-line BR treatment (no <i>TP53</i> mutation)	Dead from disease	0.0027	There is limited good quality data from Phase III clinical trials on first-line BR treatment efficacy in the published literature. That which does exist [323-325] suggests that clinical outcomes will be similar to those achieved by FCR treatment (including the impact of del17p status). Consequently, the TPs for first-line BR treatment were assumed to be the same as those for first-line FCR treatment.	[230,323-325]
p5	Undergoing first-line BR treatment (no <i>TP53</i> mutation)	No response to treatment – enters second-line treatment	0.0016	See p4.	[230]
p6	Undergoing first-line BR treatment (no <i>TP53</i> mutation)	High risk genetic mutation emerges – enters refractory therapy	0.0006	See p3.	[321,322]

P	Disease state	Transition	TP	Notes	Source
p7	Undergoing second-line FCR treatment (no <i>TP53</i> mutation)	Dead from disease	0.0033	Robak <i>et al.</i> [302] (a Phase III clinical trial) reports a mortality rate for this patient subgroup at a maximum follow-up of 1729 days of 0.23 (62/274). However, this information is not presented by del17p status. Badoux <i>et al.</i> [303] do report such information (median overall survival is 10.5 months in del17p patients and 46.5 months in all patients), but this is a Phase II study. This hazard rate was therefore applied to the mortality rate for all patients in Robak <i>et al.</i> to generate a mortality rate for del17p patients at 1729 days. From this, the number of patients in Robak <i>et al.</i> without del17p who died was calculated, along with a mortality rate (and 28-day TP) for this patient subgroup.	[302,303]
p8	Undergoing second-line FCR treatment (no <i>TP53</i> mutation)	No response to treatment – enters refractory therapy	0.0049	Robak <i>et al.</i> [302] reports that the non-response rate in all patients receiving FCR as second-line treatment was 0.3007 at maximum follow-up, yielding a 28-day TP of 0.0058. However this information is again not presented by del17p status. Instead, data from Badoux <i>et al.</i> [303] was used to calculate a hazard rate which was applied to the response rate for all patients in Robak <i>et al.</i> to generate a response rate for del17p patients. From this, the number of patients in Robak <i>et al.</i> who did not respond was calculated, along with a 28-day TP for this patient subgroup. This 28-day TP incorporates both patients who enter BSC and patients who go on to receive refractory treatment. There is limited information available in the published literature to separate out these two groups. However, it is clear that few patients receive just two courses of FCR or BR before entering BSC [326,327]. It was therefore assumed that most of these patients (95%) received refractory treatment, with the remaining 5% entering BSC. The 28-day TP was therefore adjusted accordingly.	[302,303]
p9	Undergoing second-line FCR treatment (no <i>TP53</i> mutation)	No response to treatment – enters BSC	0.0003	See p8.	[302,303]
p10	Undergoing second-line BR treatment (no <i>TP53</i> mutation)	Dead from disease	0.0178	Median survival is reported for various patient subgroups, but not the subgroup of patients without del17p. Therefore a weighted average was taken of median survival in three other patient subgroups (trisomy 12, 13q deletion, no abnormalities) and used to calculate a 28-day TP for the patient subgroup without del17p.	[328]

P	Disease state	Transition	TP	Notes	Source
p11	Undergoing second-line BR treatment (no <i>TP53</i> mutation)	No response to treatment – enters refractory therapy	0.0049	There is limited good quality data from Phase III clinical trials on second-line BR treatment response rates in the published literature. Therefore the same assumption was made as for first-line BR treatment, namely that the second-line BR treatment response rate is similar to that achieved by second-line FCR treatment (including the impact of del17p status). Consequently, the TPs for response to second-line BR treatment were assumed to be the same as those for second-line FCR treatment.	[302,303]
p12	Undergoing second-line BR treatment (no <i>TP53</i> mutation)	No response to treatment – enters BSC	0.0003	See p11.	[302,303]
p13	Undergoing refractory Ofa treatment (no <i>TP53</i> mutation)	Dead from disease	0.0288	Byrd <i>et al.</i> [219] reports a 12 month mortality rate of 0.19 for all patients receiving refractory ofatumumab treatment. Wierda <i>et al.</i> [329] reports 18 month mortality rates of 0.59 (fludarabine- and alemtuzumab-refractory patients) and 0.54 (bulky fludarabine-refractory patients). In the absence of any better quality data in the published literature for this parameter, an unweighted average TP was calculated across both data sources. This TP was assumed to apply to all patients receiving refractory ofatumumab treatment, regardless of del17p status.	[219,329]
p14	Undergoing refractory Ofa treatment (no <i>TP53</i> mutation)	No response to treatment – enters BSC	0.0201	The non-response rate in this patient subgroup was reported to be 0.41 (42/102) at two years. A 28-day TP was calculated based on this figure.	[329]
p15	Undergoing refractory Ofa treatment (no <i>TP53</i> mutation)	Undergoing BMT	0.0174	There was no data available in the published literature to inform the estimation of this parameter. Clinicians at OUH estimated that 10% of patients treated with ofatumumab would be fit enough to be offered a BMT. This estimate was converted into a 28-day TP, and assumed to apply to all patients, regardless of del17p status.	EO
p16	Undergoing BMT (no <i>TP53</i> mutation)	Dead from disease	0.0082	Mortality rate at four years reported to be 0.35 in all patients (unaffected by del17p status). A 28-day TP was calculated based on this figure.	[330]
p17	Undergoing BMT (no <i>TP53</i> mutation)	Enters BSC	0.0097	Relapse incidence at four years reported to be 0.4 (unaffected by del17p status). It was assumed that this represented a reasonable proxy for entering BSC following BMT. A 28-day TP was therefore calculated based on this figure.	[330]

P	Disease state	Transition	TP	Notes	Source
p18	In remission following first-line treatment with FCR/BR (no <i>TP53</i> mutation)	Relapse – enters second-line treatment	0.0084	Median time to new treatment is reported to be 68.4 months. This information was used to calculate a 28-day TP for all patients (regardless of del17p status). Data on time to new treatment is not presented by del17p status, however data on PFS by del17p status is presented. 28-day TPs were therefore calculated for PFS by del17p status and the ratio of these TPs was used to impute a 28-day TP for the non-response rate in del17p patients. This TP was assumed to apply to patients receiving either FCR or BR.	[230]
p19	In remission following second-line treatment with FCR/BR (no <i>TP53</i> mutation)	Relapse – enters refractory treatment	0.0185	The proportion of all patients who had progressed at 49 months was reported by Wierda <i>et al.</i> [331] (see p53). No information was provided on progression rates by del17p status. However, Robak <i>et al.</i> [302] and Badoux <i>et al.</i> [303] provide information on overall survival and treatment response rates which can be used to calculate 28-day TPs for these two variables by del17p status (see p7, p8 and p9 for further details). Hazard rates were calculated for no del17p patients versus all patients for each of these 28-day TPs, an unweighted average was calculated, and this was applied to p53 to generate this TP. This TP was assumed to apply to patients receiving either FCR or BR.	[302,303,331]
p20	In remission following refractory treatment with Ofa (no <i>TP53</i> mutation)	Relapse – enters BSC	0.0737	The median time to new treatment is reported to be 8.5 months (fludarabine- and alemtuzumab-refractory patients) and 8.2 months (bulky fludarabine-refractory patients). A weighted average of the 28-day TPs derived from this information was calculated. As no differences by del17p status were noted, this TP was used for all patients regardless of del17p status.	[329]
p21	Receiving best-supportive care (no <i>TP53</i> mutation)	Dead from disease	0.0753	The NICE submission for the use of ofatumumab in patients refractory to fludarabine and alemtuzumab also reported survival data for patients receiving best supportive care. This information was presented as a Weibull survival function, which was used to calculate a 28-day TP. In the absence of any data to the contrary in the published literature, this probability was assumed to apply to all patients receiving BSC, regardless of previous treatments of del17p status.	[197]
p22	Undergoing first-line Ofa treatment (<i>TP53</i> mutation)	Dead from disease	0.0288	As there is no trial data available for Ofa monotherapy in treatment naïve patients, the same outcomes were assumed as for relapsed/refractory patients. See p13 for further details.	[219,329]
p23	Undergoing first-line Ofa treatment (<i>TP53</i> mutation)	No response to treatment – enters refractory therapy	0.0440	As there is no trial data available for Ofa monotherapy in treatment naïve patients, the same outcomes were assumed as for relapsed/refractory patients. See p29 for further details. As per p8, it was assumed that a small proportion of these patients will enter BSC rather than refractory treatment, hence p29 is adjusted accordingly.	[329]

P	Disease state	Transition	TP	Notes	Source
p24	Undergoing first-line Ofa treatment (<i>TP53</i> mutation)	Undergoing BMT	0.0174	See p15.	EO
p25	Undergoing first-line Ofa treatment (<i>TP53</i> mutation)	No response to treatment – enters BSC	0.0023	As there is no trial data available for Ofa monotherapy in treatment naïve patients, the same outcomes were assumed as for relapsed/refractory patients. See p29 for further details. As per p8, it was assumed that a small proportion of these patients will enter BSC rather than refractory treatment, hence p29 is adjusted accordingly.	[329]
p26	Undergoing BMT (<i>TP53</i> mutation)	Dead from disease	0.0082	See p16.	[330]
p27	Undergoing BMT (<i>TP53</i> mutation)	Enters BSC	0.0097	See p17.	[330]
p28	Undergoing refractory Ofa treatment (<i>TP53</i> mutation)	Dead from disease	0.0288	See p13.	[219,329]
p29	Undergoing refractory Ofa treatment (<i>TP53</i> mutation)	No response to treatment – enters BSC	0.0463	The non-response rate in this patient subgroup was reported to be 0.71 (22/31) at two years. A 28-day TP was calculated based on this figure.	[329]
p30	In remission following first-line treatment with Ofa (<i>TP53</i> mutation)	Relapse – enters refractory treatment	0.0737	As there is no trial data available for Ofa monotherapy in treatment naïve patients, the same outcomes were assumed as for relapsed/refractory patients. See p31 for further details.	[329]
p31	In remission following refractory treatment with Ofa (<i>TP53</i> mutation)	Relapse – enters BSC	0.0737	See p20.	[329]
p32	Receiving best-supportive care (<i>TP53</i> mutation)	Dead from disease	0.0753	See p21.	[197]
p33	Undergoing refractory IB treatment (no <i>TP53</i> mutation)	Dead from disease	0.0032	The mortality rate for this patient subgroup at 30 months was reported to be 0.10. This was converted into a 28-day TP.	[227]
p34	Undergoing refractory IB treatment (no <i>TP53</i> mutation)	No response to treatment – enters BSC	0.0043	This TP was assumed to be equal to the 28-day TP of progression whilst undergoing IB treatment. Byrd <i>et al.</i> [227] reports that this rate of progression was 0.13 at 30 months for this patient subgroup. This was converted into a 28-day TP.	[227]
p35	Undergoing refractory IB treatment (<i>TP53</i> mutation)	Dead from disease	0.0131	The mortality rate for this patient subgroup at 30 months was reported to be 0.35. This was converted into a 28-day TP.	[227]
p36	Undergoing refractory IB treatment (<i>TP53</i> mutation)	No response to treatment – enters BSC	0.0223	This TP was assumed to be equal to the 28-day TP of progression whilst undergoing IB treatment. Byrd <i>et al.</i> [227] reports that this rate of progression was 0.52 at 30 months for this patient subgroup. This was converted into a 28-day TP.	[227]

P	Disease state	Transition	TP	Notes	Source
p37	Undergoing first-line FCR treatment (unknown <i>TP53</i> mutation status)	Dead from disease	0.0035	The mortality rate for this patient subgroup was reported to be 0.13 (53/408) at three years. This was converted into a 28-day TP.	[230]
p38	Undergoing first-line FCR treatment (unknown <i>TP53</i> mutation status)	No response to treatment – enters second-line treatment	0.0026	The non-response rate for this patient subgroup was reported to be 0.10 (39/408) at three years. This was converted to a 28-day TP.	[230]
p39	Undergoing first-line BR treatment (unknown <i>TP53</i> mutation status)	Dead from disease	0.0035	See p4.	[230]
p40	Undergoing first-line BR treatment (unknown <i>TP53</i> mutation status)	No response to treatment – enters second-line treatment	0.0026	See p4.	[230]
p41	Undergoing second-line FCR treatment (unknown <i>TP53</i> mutation status)	Dead from disease	0.0041	The mortality rate after 1729 days was reported to be 0.23 (62/274). This was converted to a 28-day TP.	[302]
p42	Undergoing second-line FCR treatment (unknown <i>TP53</i> mutation status)	No response to treatment – enters refractory therapy	0.0055	See p8.	[302,303]
p43	Undergoing second-line FCR treatment (unknown <i>TP53</i> mutation status)	No response to treatment – enters BSC	0.0003	See p8.	[302,303]
p44	Undergoing second-line BR treatment (unknown <i>TP53</i> mutation status)	Dead from disease	0.0186	Median survival for all patients was reported to be 33.9 months. This was converted into a 28-day TP.	[328]
p45	Undergoing second-line BR treatment (unknown <i>TP53</i> mutation status)	No response to treatment – enters refractory therapy	0.0055	See p11.	[302,303]
p46	Undergoing second-line BR treatment (unknown <i>TP53</i> mutation status)	No response to treatment – enters BSC	0.0003	See p11.	[302,303]

P	Disease state	Transition	TP	Notes	Source
p47	Undergoing refractory Ofa treatment (unknown <i>TP53</i> mutation status)	Dead from disease	0.0288	See p13.	[219,329]
p48	Undergoing refractory Ofa treatment (unknown <i>TP53</i> mutation status)	No response to treatment – enters BSC	0.0250	Wierda <i>et al.</i> [329] reports two-yearly non-response rates of 0.42 (fludarabine- and alemtuzumab-refractory patients) and 0.53 (bulky fludarabine-refractory patients). A weighted average of these non-response rates was used to calculate a 28-day TP.	[329]
p49	Undergoing refractory Ofa treatment (unknown <i>TP53</i> mutation status)	Undergoing BMT	0.1000	See p15.	EO
p50	Undergoing BMT (unknown <i>TP53</i> mutation status)	Dead from disease	0.0082	See p16.	[330]
p51	Undergoing BMT (unknown <i>TP53</i> mutation status)	Enters BSC	0.0097	See p17.	[330]
p52	In remission following first-line treatment with FCR/BR (unknown <i>TP53</i> mutation status)	Relapse – enters second-line treatment	0.0093	Median time to new treatment is reported to be 68.4 months. This information was used to calculate a 28-day TP for all patients (regardless of del17p status). This TP was assumed to apply to patients receiving either FCR or BR.	[230]
p53	In remission following second-line treatment with FCR/BR (unknown <i>TP53</i> mutation status)	Relapse – enters refractory treatment	0.0218	The proportion of all patients who had progressed at 49 months was reported to be 0.69. This figure was converted into a 28-day TP. This TP was assumed to apply to patients receiving either FCR or BR.	[331]
p54	In remission following refractory treatment with Ofa (unknown <i>TP53</i> mutation status)	Relapse – enters BSC	0.0737	See p20.	[329]
p55	Receiving best-supportive care (unknown <i>TP53</i> mutation status)	Dead from disease	0.0753	See p21.	[197]
p56	Undergoing first-line FCR treatment (patient correctly identified by genomic testing to be an FCR responder)	Dead from disease	0.0022	It was assumed that the efficacy of FCR treatment would be improved in this patient subgroup due to better targeting of treatment. The mortality rate at three years estimated by Hallek <i>et al.</i> [230] was therefore applied for patients without del17p, reduced by 20% to 0.08. This was converted to a 28-day TP.	[230]

P	Disease state	Transition	TP	Notes	Source
p57	Undergoing first-line FCR treatment (patient correctly identified by genomic testing to be an FCR responder)	No response to treatment – enters second-line treatment	0.0018	It was assumed that the efficacy of FCR treatment would be improved in this patient subgroup due to better targeting of treatment. The non-response rate at three years estimated by Hallek <i>et al.</i> [230] was therefore applied for patients without del17p, reduced by 20% to 0.07. This was converted to a 28-day TP.	[230]
p58	Undergoing first-line FCR treatment (patient correctly identified by genomic testing to be an FCR responder)	Further genomic changes emerge – enters refractory therapy	0.0006	See p3.	[321,322]
p59	Undergoing first-line FCR treatment (patient incorrectly identified by genomic testing to be an FCR responder)	Dead from disease	0.0244	In the absence of any data in the published literature, it was assumed that this TP was equal to that for mortality in del17p patients receiving FCR as first-line treatment. The survival rate in this patient subgroup was reported to be 0.38 (8/22). This was converted to a 28-day TP.	[230]
p60	Undergoing first-line FCR treatment (patient incorrectly identified by genomic testing to be an FCR responder)	No response to treatment – enters second-line treatment	0.0097	In the absence of any data in the published literature, it was assumed that this TP was equal to that for non-response in del17p patients receiving FCR as first-line treatment. The non-response rate in this patient subgroup was reported to be 0.32 (7/22). This was converted to a 28-day TP.	[230]
p61	Undergoing first-line FCR treatment (patient incorrectly identified by genomic testing to be an FCR responder)	Further genomic changes emerge – enters refractory therapy	0.0006	See p3.	[321,322]
p62	Undergoing second-line BR treatment (patient identified by genomic testing to be a likely FCR responder)	Dead from disease	0.0128	It was assumed that the efficacy of BR treatment would be improved in this patient subgroup due to better targeting of treatment. An estimate of median survival was derived in the same way as for p10, which was then increased by 20%. This was converted to a 28-day TP.	[328]
p63	Undergoing second-line BR treatment (patient identified by genomic testing to be a likely FCR responder)	No response to treatment – enters refractory therapy	0.0028	It was assumed that the efficacy of BR treatment would be improved in this patient subgroup due to better targeting of treatment. An estimate of the non-response rate was derived in the same way as for p8 and p9. This was then decreased by 20%. This was converted to a 28-day TP.	[302,303]
p64	Undergoing second-line BR treatment (patient identified by genomic testing to be a likely FCR responder)	No response to treatment – enters BSC	0.0012	See p63.	[302,303]

P	Disease state	Transition	TP	Notes	Source
p65	Undergoing refractory IB treatment (patient identified by genomic testing to be a likely FCR responder)	Dead from disease	0.0032	In the absence of any data in the published literature, it was assumed that the mortality rate in patients without del17p was a reasonable proxy for the mortality rate in likely FCR responders. See p33 for further information.	[227]
p66	Undergoing refractory IB treatment (patient identified by genomic testing to be a likely FCR responder)	No response to treatment – enters BSC	0.0043	In the absence of any data in the published literature, it was assumed that the non-response rate in patients without del17p was a reasonable proxy for the mortality rate in likely FCR responders. See p34 for further information.	[227]
p67	In remission following first-line treatment with FCR (patient correctly identified by genomic testing to be an FCR responder)	Relapse – enters second-line treatment	0.0070	It was assumed that the efficacy of FCR treatment would be improved in this patient subgroup due to better targeting of treatment. The median time to new treatment estimated by Hallek <i>et al.</i> [230] was therefore applied for patients without del17p, increased by 20% to 2754 days. This was converted to a 28-day TP.	[230]
p68	In remission following first-line treatment with FCR (patient incorrectly identified by genomic testing to be an FCR responder)	Relapse – enters second-line treatment	0.0361	In the absence of any data in the published literature, it was assumed that this TP was equal to that for time to new treatment following FCR in del17p patients. See p18 for further details.	[230]
p69	In remission following second-line treatment with BR (patient identified by genomic testing to be a likely FCR responder)	Relapse – enters refractory treatment	0.0127	It was assumed that the efficacy of BR treatment would be improved in this patient subgroup. The rate of progression used in p53 was therefore applied, reduced by 20% to 0.55. This was then converted to a 28-day TP.	[331]
p70	Receiving best-supportive care (patient identified by genomic testing to be a likely FCR responder)	Dead from disease	0.0753	See p21.	[197]
p71	Undergoing first-line IB treatment (patient identified by genomic testing to be a likely FCR non-responder)	Dead from disease	0.0025	There is very limited data in the published literature on the efficacy of first-line IB treatment. O'Brien <i>et al.</i> [229] reports such data for a small group of elderly patients (who may be a reasonable proxy for likely FCR non-responders) estimating that the proportion of patients who are dead at 24 months is 0.034. Farooqui <i>et al.</i> [228] reports such data for a small group of patients with del17p, estimating that the proportion of patients who are dead at 24 months is 0.09. Given the limitations associated with the two datasets, 28-day TPs were calculated for both and an unweighted average of these two TPs was used.	[228,229]

P	Disease state	Transition	TP	Notes	Source
p72	Undergoing first-line IB treatment (patient identified by genomic testing to be a likely FCR non-responder)	No response to treatment – enters BSC	0.0041	This TP was estimated using the same studies as p71. The progression rate at 24 months is reported to be 0.037 (O'Brien <i>et al.</i> [229]) and 0.16 (Farooqui <i>et al.</i> [228]). 28-day TPs were calculated for both studies and an unweighted average used.	[228,229]
p73	Receiving best-supportive care (patient identified by genomic testing to be a likely FCR non-responder)	Dead from disease	0.0753	See p21.	[197]

BMT = bone marrow transplant; BR = bendamustine and rituximab; BSC = best supportive care; EO = expert opinion; FCR = fludarabine, cyclophosphamide and rituximab; IB = ibrutinib; NICE = National Institute for Health and Care Excellence; Ofa = ofatumumab; OUH = Oxford University Hospitals NHS Trust; P = Probability; PFS = progression-free survival; TP = 28-day transition probability.

Appendix 15: Resource use and overall costs for each health state

State	Item ^a	Notes ^b	Cost (£)
FCR treatment (first-line, first cycle)	Genetic test (Comparator A/B only)	-	281.08
	Genomic test (Intervention 1/2 only)	-	319.23
	FCR treatment	-	2554.74
	Co-trimoxazole treatment	960mg three times a week	2.29
	Fluconazole treatment	50mg daily	0.76
	Aciclovir treatment	200mg three times a day	2.59
	GCSF treatment	60% of patients receive 300mcg daily for seven days of cycle	234.26
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan per cycle	106.32
	Blood transfusion	20% of patients have one transfusion over all cycles of FCR treatment. Average treatment length is 5.2 cycles, so 0.19 transfusions per cycle were costed	14.69
	Outpatient appointment with haematologist	One appointment per cycle	142.71
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications (Comparator A/B/C)	See Appendix 18 for more details of complication rates	968.83
	Complications (Intervention 1/2)	See Appendix 18 for more details of complication rates	484.42
		TOTAL (Comparator A/B)	4404.48
		TOTAL (Comparator C)	4123.40
		TOTAL (Intervention 1/2)	3958.22

State	Item ^a	Notes ^b	Cost (£)
FCR treatment (second-line, first cycle)	Genetic test (Comparator A/B only)		281.08
	FCR treatment	-	2554.74
	Co-trimoxazole treatment	960mg three times a week	2.29
	Fluconazole treatment	50mg daily	0.76
	Aciclovir treatment	200mg three times a day	2.59
	GCSF treatment	60% of patients receive 300mcg daily for seven days of cycle	234.26
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan per cycle	106.32
	Blood transfusion	20% of patients have one transfusion over all cycles of FCR treatment. Average treatment length is 5.2 cycles, so 0.19 transfusions per cycle were costed	14.69
	Outpatient appointment with haematologist	One appointment per cycle	142.71
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications	See Appendix 18 for more details of complication rates	1937.66
		TOTAL (Comparator A/B)	5373.31
		TOTAL (Comparator C)	5092.23

State	Item ^a	Notes ^b	Cost (£)
FCR treatment (first- or second-line, subsequent cycles)	FCR treatment	-	2876.02
	Co-trimoxazole treatment	960mg three times a week	2.29
	Fluconazole treatment	50mg daily	0.76
	Aciclovir treatment	200mg three times a day	2.59
	GCSF treatment	60% of patients receive 300mcg daily for seven days of cycle	234.26
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan is given at the end of the full course of FCR treatment. Average treatment length is 4.2 cycles (excluding cycle 1), so 0.24 scans per cycle were costed	25.31
	Blood transfusion	20% of patients have one transfusion over all cycles of FCR treatment. Average treatment length is 5.2 cycles, so 0.19 transfusions per cycle were costed	14.69
	Outpatient appointment with haematologist	After cycle one, patients have an outpatient appointment halfway through the treatment course and another at the end. Average treatment length is 4.2 cycles (excluding cycle one), so 0.48 appointments per cycle were costed	67.96
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications (Comparator A/B/C, first-line treatment)	See Appendix 18 for more details of complication rates	968.83
	Complications (Comparator A/B/C, second-line treatment)	See Appendix 18 for more details of complication rates	1937.66
	Complications (Intervention 1/2)	See Appendix 18 for more details of complication rates	484.42
TOTAL (Comparator A/B/C, first-line)			4288.92
TOTAL (Comparator A/B/C, second-line)			5257.75
TOTAL (Intervention 1/2)			3804.50

State	Item ^a	Notes ^b	Cost (£)
BR treatment (first-line, first cycle)	Genetic test (Comparator A/B only)	-	281.08
	BR treatment	-	2854.63
	GCSF treatment	10% of patients receive 300mcg daily for seven days of cycle	39.04
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan per cycle	106.32
	Blood transfusion	10% of patients have one transfusion over all cycles of BR treatment. Average treatment length is 4.5 cycles, so 0.22 transfusions per cycle were costed	8.44
	Outpatient appointment with haematologist	One appointment per cycle	142.71
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications	See Appendix 18 for more details of complication rates	1006.85
		TOTAL (Comparator A/B)	4535.29
	TOTAL (Comparator C)	4254.21	
BR treatment (second-line, first cycle)	Genetic test (Comparator A/B only)	-	281.08
	Genomic test (Intervention 1/2 only)	-	319.23
	BR treatment	-	2854.63
	GCSF treatment	10% of patients receive 300mcg daily for seven days of cycle	39.04
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan per cycle	106.32
	Blood transfusion	10% of patients have one transfusion over all cycles of BR treatment. Average treatment length is 4.5 cycles, so 0.22 transfusions per cycle were costed	8.44
	Outpatient appointment with haematologist	One appointment per cycle	142.71
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications (Comparator A/B/C)	See Appendix 18 for more details of complication rates	2013.69
	Complications (Intervention 1/2)	See Appendix 18 for more details of complication rates	1006.85
		TOTAL (Comparator A/B)	5542.13
		TOTAL (Comparator C)	5261.05
	TOTAL (Intervention 1/2)	4573.44	

State	Item ^a	Notes ^b	Cost (£)
BR treatment (first- or second-line, subsequent cycles)	BR treatment	-	2854.63
	GCSF treatment	10% of patients receive 300mcg daily for seven days of cycle	39.04
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan is given at the end of the full course of FCR treatment. Average treatment length is 3.5 cycles (excluding cycle 1), so 0.28 scans per cycle were costed	30.16
	Blood transfusion	10% of patients have one transfusion over all cycles of BR treatment. Average treatment length is 4.5 cycles, so 0.22 transfusions per cycle were costed	8.44
	Outpatient appointment with haematologist	After cycle one, patients have an outpatient appointment halfway through the treatment course and another at the end. Average treatment length is 3.5 cycles (excluding cycle one), so 0.57 appointments per cycle were costed	80.95
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications (Comparator A/B/C, first-line treatment)	See Appendix 18 for more details of complication rates	1006.85
	Complications (Comparator A/B/C, second-line treatment)	See Appendix 18 for more details of complication rates	2013.69
	Complications (Intervention 1/2)	See Appendix 18 for more details of complication rates	1006.85
		TOTAL (Comparator A/B/C, first-line)	4427.32
		TOTAL (Comparator A/B/C, second-line)	5434.16
		TOTAL (Intervention 1/2)	4427.32

State	Item ^a	Notes ^b	Cost (£)
Ofa treatment (first-line, first cycle)	Genetic test (Comparator A/B only)	-	281.08
	Genomic test (Intervention 2 only)	-	319.23
	Ofa treatment	-	11876.84
	Hydrocortisone treatment	100mg daily for 28 days	30.24
	Chlorpheniramine treatment	30mg daily for 28 days	1.65
	Paracetamol	1g daily for 28 days	0.32
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan per cycle	106.32
	Blood transfusion	10% of patients have one transfusion over all cycles of Ofa treatment. Average treatment length is 5.8 cycles, so 0.17 transfusions per cycle were costed	6.63
	Outpatient appointment with haematologist	One appointment per cycle	142.71
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications	See Appendix 18 for more details of complication rates	255.93
TOTAL (Comparator A/B)			12797.93
TOTAL (Intervention 2)			12836.08
Ofa treatment (first-line, second cycle)	Ofa treatment	-	14942.86
	Hydrocortisone treatment	100mg daily for 28 days	30.24
	Chlorpheniramine treatment	30mg daily for 28 days	1.65
	Paracetamol	1g daily for 28 days	0.32
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	Blood transfusion	10% of patients have one transfusion over all cycles of Ofa treatment. Average treatment length is 5.8 cycles, so 0.17 transfusions per cycle were costed	6.63
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications	See Appendix 18 for more details of complication rates	255.93
TOTAL			15333.84

State	Item ^a	Notes ^b	Cost (£)
Ofa treatment (first-line, subsequent cycles)	Ofa treatment	-	3992.11
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan is given at the end of the full course of Ofa treatment. Average treatment length is 3.8 cycles (excluding cycles one and two), so 0.27 scans per cycle were costed	28.27
	Blood transfusion	10% of patients have one transfusion over all cycles of Ofa treatment. Average treatment length is 5.8 cycles, so 0.17 transfusions per cycle were costed	6.63
	Outpatient appointment with haematologist	After cycles one and two, patients have an outpatient appointment halfway through the treatment course and another at the end. Average treatment length is 3.8 cycles (excluding cycles one and two), so 0.53 appointments per cycle were costed	75.88
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications	See Appendix 18 for more details of complication rates	255.93
		TOTAL	4455.03

State	Item ^a	Notes ^b	Cost (£)
Ofa treatment (second-line or refractory, first cycle)	Ofa treatment	-	11876.84
	Genetic test (Comparator A/B only)	-	281.08
	Hydrocortisone treatment	100mg daily for 28 days	30.24
	Chlorpheniramine treatment	30mg daily for 28 days	1.65
	Paracetamol	1g daily for 28 days	0.32
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan per cycle	106.32
	Blood transfusion	10% of patients have one transfusion over all cycles of Ofa treatment. Average treatment length is 5.8 cycles, so 0.17 transfusions per cycle were costed	6.63
	Outpatient appointment with haematologist	One appointment per cycle	142.71
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications	See Appendix 18 for more details of complication rates	511.87
TOTAL (Comparator A/B)			13053.87
TOTAL (Comparator C)			12772.79
Ofa treatment (second-line or refractory, second cycle)	Ofa treatment	-	14942.86
	Hydrocortisone treatment	100mg daily for 28 days	30.24
	Chlorpheniramine treatment	30mg daily for 28 days	1.65
	Paracetamol	1g daily for 28 days	0.32
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	Blood transfusion	10% of patients have one transfusion over all cycles of Ofa treatment. Average treatment length is 5.8 cycles, so 0.17 transfusions per cycle were costed	6.63
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications	See Appendix 18 for more details of complication rates	511.87
TOTAL			15589.78

State	Item ^a	Notes ^b	Cost (£)
Ofa treatment (second-line or refractory, subsequent cycles)	Ofa treatment	-	3992.11
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan is given at the end of the full course of Ofa treatment. Average treatment length is 3.8 cycles (excluding cycles one and two), so 0.27 scans per cycle were costed	28.27
	Blood transfusion	10% of patients have one transfusion over all cycles of Ofa treatment. Average treatment length is 5.8 cycles, so 0.17 transfusions per cycle were costed	6.63
	Outpatient appointment with haematologist	After cycles one and two, patients have an outpatient appointment halfway through the treatment course and another at the end. Average treatment length is 3.8 cycles (excluding cycles one and two), so 0.53 appointments per cycle were costed	75.88
	Registrar appointment prior to treatment commencing	One appointment per cycle	90.72
	Complications	See Appendix 18 for more details of complication rates	511.87
		TOTAL	4710.97
Ibrutinib treatment (first-line)	Genomic test	Patients undergo one test across the full course of IB treatment. Average treatment length is 9.3 cycles, so 0.10 tests per cycle were costed	34.15
	IB treatment	-	2472.10
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan is given at the start of treatment, and another after three months. Average treatment length is 9.3 cycles, so 0.21 scans per cycle were costed	22.75
	IgV _H test	One test per cycle	38.74
	Blood transfusion	10% of patients have one transfusion over all cycles of IB treatment. Average treatment length is 9.3 cycles, so 0.11 transfusions per cycle were costed	4.08
	Outpatient appointment with haematologist	Patients have three appointments across the whole treatment period. Average treatment length is 9.3 cycles, so 0.32 appointments per cycle were costed	45.80
	Complications	See Appendix 18 for more details of complication rates	80.73
		TOTAL	2703.84

State	Item ^a	Notes ^b	Cost (£)
Ibrutinib treatment (second-line or refractory)	Genetic test (Comparator B only)	Patients undergo one test across the full course of IB treatment. Average treatment length is 9.3 cycles, so 0.10 tests per cycle were costed	30.07
	Genomic test (Intervention 1/2 only)	As above - 0.10 tests per cycle were costed	34.15
	IB treatment	-	2472.10
	FBC	One test per cycle	3.01
	U&E	One test per cycle	1.25
	LFT	One test per cycle	1.25
	CT scan	One scan is given at the start of treatment, and another after three months. Average treatment length is 9.3 cycles, so 0.21 scans per cycle were costed	22.75
	IgV _H test	One test per cycle	38.74
	Blood transfusion	10% of patients have one transfusion over all cycles of IB treatment. Average treatment length is 9.3 cycles, so 0.11 transfusions per cycle were costed	4.08
	Outpatient appointment with haematologist	Patients have three appointments across the whole treatment period. Average treatment length is 9.3 cycles, so 0.32 appointments per cycle were costed	45.80
	Complications (Comparator B only)	See Appendix 18 for more details of complication rates	161.45
	Complications (Intervention 1/2 only)	See Appendix 18 for more details of complication rates	80.73
		TOTAL (Comparator B)	2780.49
		TOTAL (Intervention 1/2)	2703.84
Remission following FCR treatment (all lines of treatment)	Co-trimoxazole treatment	960mg three times a week. Given for three months (12 weeks) following last cycle of FCR treatment. 36 doses divided by the average number of 28 day cycles spent in remission following FCR treatment (56.3 cycles) so 0.64 doses per cycle	0.12
	Fluconazole treatment	50mg daily. Given for three months (12 weeks) following last cycle of FCR treatment. 84 doses divided by the average number of 28 day cycles spent in remission following FCR treatment (56.3 cycles) so 0.64 doses per cycle	0.04
	Aciclovir treatment	200mg three times a day. Given for three months (12 weeks) following last cycle of FCR treatment. 252 doses divided by the average number of 28 day cycles spent in remission following FCR treatment (56.3 cycles) so 4.48 doses per cycle	0.14
	FBC	One test every three cycles	1.00
	Outpatient appointment with haematologist	One appointment every three cycles	47.57
	IVIG treatment	One infusion every three weeks in 10% of patients	160.16
	Outpatient appointment with immunologist	One appointment every six months in 10% of patients	3.25
		TOTAL	212.29

State	Item ^a	Notes ^b	Cost (£)
Remission following BR treatment (all lines of treatment)	FBC	One test every three cycles	1.00
	Outpatient appointment with haematologist	One appointment every three cycles	47.57
	IVIG treatment	One infusion every three weeks in 5% of patients	80.08
	Outpatient appointment with immunologist	One appointment every six months in 5% of patients	1.63
		TOTAL	130.28
Remission following Ofa treatment (all lines of treatment)	FBC	One test every three cycles	1.00
	Outpatient appointment with haematologist	One appointment every three cycles	47.57
		TOTAL	48.57
Undergoing BMT	BMT	One per cycle	43723.96

State	Item ^a	Notes ^b	Cost (£)
Remission following BMT	Outpatient appointment with haematologist	1-2 appointments a week for the first eight weeks, then two appointments in the next four weeks, then one appointment every three weeks until 26 weeks. Assumed to be 19 appointments in total, spread over the average number of cycles spent in remission following BMT (51.3 cycles), so 0.37 appointments per cycle in total	52.83
	Outpatient appointment with haematology nurse	As above - 0.37 appointments per cycle in total	33.58
	Telephone consultation with haematology nurse	One appointment per week	60.77
	CT scan	One scan is given after three months. Average treatment length is 51.3 cycles, so 0.02 scans per cycle were costed	2.07
	Lung function test	4-6 tests required over course of entire remission period (51.3 cycles) so costed five tests at 0.10 per cycle	3.96
	Diagnostic bone marrow extraction	Three tests required over course of entire remission period (51.3 cycles) so 0.06 per cycle	28.20
	IgV _H test	Three tests required over course of entire remission period (51.3 cycles) so 0.06 per cycle	2.26
	Thyroid function test	Five tests required over course of entire remission period (51.3 cycles) so 0.10 per cycle	0.12
	Immunoglobulin test	6-7 tests required over course of entire remission period (51.3 cycles) so costed 6.5 tests at 0.13 per cycle	0.38
	IVIG treatment	5% of patients require one course of treatment, spread over the entire remission period (51.3 cycles) so 0.02 courses per cycle in total	1.17
	FBC	Six tests required over course of entire remission period (51.3 cycles) so 0.12 per cycle	0.35
	Donor lymphocyte infusion	25% of patients require infusions. Marks <i>et al.</i> [332] report that 130 infusions were required across 81 patients. This figure was divided by the average number of cycles spent in remission following BMT, hence 0.03 infusions per cycle	2.99
	Blood and platelet support	15% of patients require transfusions, with 14 transfusions required on average over the three month period following BMT. This figure was divided by the average number of cycles spent in remission following BMT hence 0.27 transfusions per cycle	15.63
	Complications	See Appendix 18 for more details of complication rates	28.72
TOTAL			233.04

State	Item ^a	Notes ^b	Cost (£)
BSC	Fluconazole treatment	50mg daily in 20% of patients	0.15
	Aciclovir treatment	200mg three times a day in 20% of patients	0.52
	IVIG treatment	One infusion every three weeks in 20% of patients	320.33
	FBC	One test per cycle	3.01
	Outpatient appointment with haematologist	One appointment per cycle	142.71
	Blood transfusion	One transfusion every five weeks in 50% of patients	152.75
	Inpatient stay	One appointment every 2-3 months (assumed to be every 2.5 months)	1030.86
TOTAL			1650.33

BMT = bone marrow transplant; BR = bendamustine and rituximab; BSC = best supportive care; CT = computerised tomography; FBC = full blood count; FCR = fludarabine, cyclophosphamide and rituximab; GCSF = granulocyte colony-stimulating factor; IB = ibrutinib; IgV_H = immunoglobulin heavy chain variable region; IVIG = IV immunoglobulin treatment; LFT = liver function test; mcg = micrograms; mg = milligrams; Ofa = ofatumumab; U&E = urea and electrolytes test. ^a All chemotherapy costs include the cost of chemotherapy drugs, pharmacy costs and also the costs of treatment delivery. ^b A number of the calculations in this table incorporate information on the average number of cycles of each type of chemotherapy treatment. For further information on these calculations, please see **Appendix 18**.

Appendix 16: Unit costs

Resource use item	Unit cost	Notes	Source
<i>Chemotherapy drugs^a</i>			
FCR (first cycle)	£2,224.87	Cost per 28-day cycle. Fludarabine: 40mg/m ² for five days of cycle. Cyclophosphamide: 250mg/m ² per day for first three days of cycle. Rituximab: 375mg/m ² on day zero of cycle	[233,234]
FCR (subsequent cycles)	£2,574.13	Cost per 28-day cycle. Fludarabine: 40mg/m ² for five days each cycle. Cyclophosphamide: 250mg/m ² per day for first three days of cycle. Rituximab: 500mg/m ² on day one of cycle	[233,234]
BR (first cycle)	£2,504.26	Cost per 28-day cycle. Bendamustine: 100mg/m ² on days one and two of cycle. Rituximab: 375mg/m ² on day zero of cycle	[234]
BR (subsequent cycles)	£2,853.52	Cost per 28-day cycle. Bendamustine: 100mg/m ² on days one and two of cycle. Rituximab: 500mg/m ² on day one of cycle	[234]
Ofatumumab (first cycle)	£11,466.00	Cost per 28-day cycle. One infusion per week (infusion one is 300mg, infusions 2-4 are 2,000mg). For ofatumumab with high dose methylprednisolone, add £105.80. For ofatumumab with dexamethasone, add £3.85	[233,234]
Ofatumumab (second cycle)	£14,560.00	Cost per 28-day cycle. One 2,000mg infusion per week. For ofatumumab with high dose methylprednisolone, add £105.80. For ofatumumab with dexamethasone, add £3.85	[233,234]
Ofatumumab (subsequent cycles)	£3,640.00	Cost per 28-day cycle. One 2,000mg infusion per cycle. For ofatumumab with high dose methylprednisolone, add £105.80. For ofatumumab with dexamethasone, add £3.85	[233,234]
Ibrutinib (all cycles)	£2,299.79	Cost per 28-day cycle. Daily dose 420mg. Assumed that annual cost of treatment would not exceed £30,000, so daily cost assumed to be £82.14	Ass
Treatment of complications arising from chemotherapy	£2,776.70	An unweighted average was taken of the costs of seven potential complications arising from chemotherapy. Further information provided in Appendix 19 .	[232,234,333]
<i>Other drugs^a</i>			
Co-trimoxazole	£0.19	Cost per 960mg dose (given three times a week)	[233]
Fluconazole	£0.03	Cost per 50mg dose (given daily)	[233]
Aciclovir	£0.03	Cost per 200mg dose (given three times a day)	[233]
GCSF	£55.78	Cost per 300mcg injection (given daily)	[233]
Hydrocortisone	£1.08	Cost per 100mg injection (given daily)	[233]
Chlorpheniramine	£0.06	Cost per 30mg dose (given daily)	[233]
Paracetamol	£0.01	Cost per 1g dose (given daily)	[233]
Ganciclovir	£29.77	Cost per 5mg/kg IV infusion (given twice daily)	[234]
IV immunoglobulin treatment	£1,201.23	Immunoglobulins, Band 1, weighted average (Currency code: XD34Z)	[232]

Resource use item	Unit cost	Notes	Source
<i>Surgeries, tests and procedures^b</i>			
Genetic test	£281.08	Calculated based on clinical practice in OUH. Further information provided in Appendix 17 .	Calc
Genomic test	£319.23	Calculated based on clinical practice in OUH. Further information provided in Appendix 17 .	Calc
FBC	£3.01	Haematology test (Currency code: DAPS05)	[232]
U&E	£1.25	Clinical biochemistry test (Currency code: DAPS04)	[232]
LFT	£1.25	Clinical biochemistry test (Currency code: DAPS04)	[232]
CT scan	£106.32	Average of all types of CT scan, weighted by activity	[232]
IgV _H test	£38.74	Histopathology and histology test (Currency code: DAPS02)	[232]
Lung function test	£40.61	Simple lung function test (Currency code: DZ32Z)	[232]
Diagnostic bone marrow extraction	£482.48	Clinical haematology (Currency code: SA33Z)	[232]
Thyroid function test	£1.25	Clinical biochemistry test (Currency code: DAPS04)	[232]
Immunoglobulin test	£3.01	Haematology test (Currency code: DAPS05)	[232]
Blood transfusion	£381.88	This cost reflects 2.10 units of CMV negative and irradiated red cells, and one haematology appointment (non-consultant)	[197,232,334]
Blood and platelet support	£381.88	This cost was assumed to be the same as that for a blood transfusion	[197,232,334]
BMT	£43,723.96	Average of all allogeneic graft categories (Currency codes SA20A, SA21A, SA22A), weighted by activity	[232]
Treatment of complications arising from BMT	£1,842.66	An unweighted average was taken of the costs of three potential complications post-BMT. Further information provided in Appendix 20 .	[232,234]
Donor lymphocyte infusion	£381.88	Assumed to be the same cost as a blood transfusion, as per a previous study in leukaemia	[197,232,334,335]
<i>Staff costs^b</i>			
Haematology appointment (consultant)	£142.71	Currency code WF01A: Non-Admitted Face to Face Attendance, Follow-up - Clinical Haematology. Service code: 303 (Consultant led outpatient attendances)	[232]
Haematology appointment (non-consultant)	£90.72	Currency code WF01A: Non-Admitted Face to Face Attendance, Follow-up - Clinical Haematology. Service code: 303 (Non consultant led outpatient attendances)	[232]
Telephone consultation with haematology nurse	£60.77	Currency code WF01C: Non-Admitted Non-Face to Face Attendance, Follow-up - Clinical Haematology. Service code: 303 (Non consultant led outpatient attendances)	[232]
Immunology appointment (consultant)	£195.01	Currency code WF01A: Non-Admitted Face to Face Attendance, Follow-up - Clinical Immunology. Service code: 316 (Consultant led outpatient attendances)	[232]

Resource use item	Unit cost	Notes	Source
<i>Miscellaneous^b</i>			
Pharmacy costs for dispensing chemotherapy drugs	£10.25	15min of hospital pharmacist time	[197,235]
Deliver exclusively oral chemotherapy	£162.06	Currency code SB11Z. Weighted average of all service codes (DCRDN, OP, Oth)	[232]
Deliver more complex parenteral chemotherapy at first attendance	£261.71	Currency code SB13Z. Weighted average of all service codes (DCRDN, OP, Oth)	[232]
Deliver complex chemotherapy, including prolonged infusional treatment, at first attendance	£319.62	Currency code SB14Z. Weighted average of all service codes (DCRDN, OP, Oth)	[232]
Deliver subsequent elements of a chemotherapy cycle	£291.64	Currency code SB15Z. Weighted average of all service codes (DCRDN, OP, Oth)	[232]
Inpatient stay	£2,577.16	Currency code SA32D: Chronic Lymphocytic Leukaemia and Related Disorders, with CC Score 0-2. Service code 303. Non-elective long stay.	[232]

Ass = assumption; BMT = bone marrow transplant; BR = bendamustine and rituximab; Calc = calculated; CMV = Cytomegalovirus; CT = computerised tomography; FBC = full blood count; FCR = fludarabine, cyclophosphamide and rituximab; GCSF = granulocyte colony-stimulating factor; IgV_H = immunoglobulin heavy chain variable region; LFT = liver function test; OUH = Oxford University Hospitals NHS Trust; U&E = urea and electrolytes test. ^a Dosages are based on an adult with height 168cm and weight 77kg (yielding a body surface area of 1.88m²) [336]. ^b All costs informed by NHS Reference Costs are based on patients aged over 18 years.

Appendix 17: Microcosting study of chronic lymphocytic leukaemia genetic and genomic testing practice at the Oxford University Hospital NHS Trust

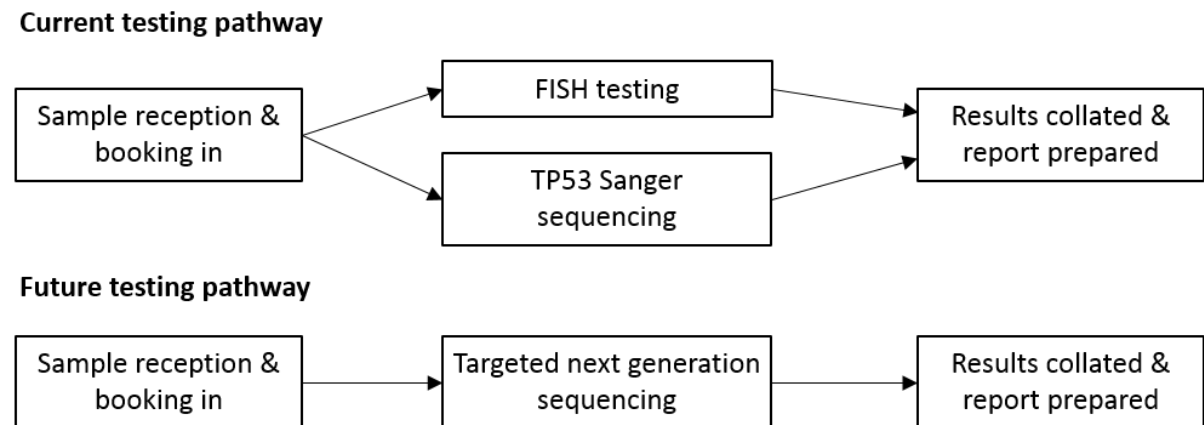
Background and methods

This microcosting study aimed to calculate the cost of pre-treatment genetic and genomic testing of chronic lymphocytic leukaemia (CLL) patient samples using:

- a) Fluorescence in situ hybridisation (FISH) testing and Sanger sequencing (Current practice)
- b) Targeted next-generation sequencing (NGS) (Future practice)

Current and likely future testing pathways are described in **Figure 1**. Current practice pathways were based on standard operating procedures (SOPs) at the Oxford Biomedical Research Centre Haemato-Molecular Diagnostic Service (HMDS), where samples are received as peripheral blood in ethylenediamine tetraacetic (EDTA) acid. The process of sample reception was assumed to be identical in both current and future pathways, but collation of results differed. Targeted NGS was assumed to take place using the Illumina MiSeq Benchtop Sequencer.

Figure 1: Current and likely future testing pathways



Information from the SOPs was used to design questionnaires to collect information on each stage of both testing pathways. These questionnaires were completed by staff at the HMDS. Where information on equipment and consumable costs was not available from staff at the HMDS, this was obtained from alternative sources e.g. laboratory equipment suppliers. In some cases, cost data was only available for a kit as a whole, rather than the individual equipment or consumable items within the kit. In these cases, kit costs were apportioned equally across all items within the kit.

Information on staff salaries was extracted from Agenda for Change salary scales for England from April 2014: mid-points of salary ranges were used. For the purposes of calculating an hourly staff rate, a working year was assumed to be 46 weeks, and a working week was assumed to be 37.5 hours. All basic staff salaries were inflated by 20% to incorporate National Insurance and Superannuation.

The HMDS was assumed to process 150 CLL samples a year. A small number of these samples were assumed to repeat specific stages of the testing process. The costs of equipment items were discounted at a rate of 3.5% over a lifespan appropriate to each equipment item. Equipment lifespans were based on information provided by laboratory staff. If no information was available, a lifespan of five years was assumed. If maintenance costs were incurred for equipment items, these were also included. Both basic equipment costs and maintenance costs were weighted by the percentage of time that a piece of equipment was used for CLL testing. Sample batching was incorporated into cost calculations. Value added tax was included in all cost estimates, calculated at 20%. Overheads were also included, calculated at 20% of the cost of testing. All costs are reported in British Pounds at 2013 values.

Results

3.1 Sample reception

Sample reception consists of two stages (**Table 1**).

Table 1: Sample reception – costs by stage

Stage	Cost category			All categories
	Equipment	Consumables	Staff	
Sample reception	£0.09	£0.02	£3.75	£3.86
Review	£0.07	£0.00	£1.64	£1.71
Total (before overheads)	£0.16	£0.02	£5.40	£5.57
% total cost	3%	0%	97%	-
TOTAL (including overheads calculated at 20%)	-	-	-	£6.69

Table 2 reports the staff time used during sample reception, the costs of this time, and cost for high and low values of staff costs. Most of the cost of sample reception is the cost of the staff time required to process samples (£4.90 of the staff costs).

Table 2: Cost of staff time for sample reception

Staff member	Time (min)	Cost per patient sample at different staff costs per hour		
		Mean cost	Low cost	High cost
Medical technical officer ^a	17.13	£4.90	£4.25	£5.54
Principle clinical scientist ^b	1.00	£0.50	£0.45	£0.55
TOTAL	18.13	£5.40	£4.70	£6.09

^a Mean cost/hour £17.14 (Low: £14.88, High: £19.41); ^b Mean cost/hour £30.03 (Low: £27.30, High: £32.76)

3.2 Fluorescence in situ hybridisation testing

FISH testing consists of six stages (**Table 3**). Most of the costs of FISH testing relate to either equipment (41%) or consumables (46%). Two stages are particularly costly. The first is probe hybridisation, which uses acetone, costing £15.60/ml. Around 40ml is used for FISH testing of CLL samples each month, costing £624 (£49.92 per sample). The second costly stage is analysis. The main cost driver here is the use of a fluorescent microscope, accounting for £46.58 of the stage costs. This costs £69,000 (plus £3,060 per year maintenance cost), lasts for five years, and 40% of the use of this item is devoted to FISH testing of CLL samples. Staff costs (**Table 4**) represent 13% of total costs: most of this cost is medical technical officer (MTO) time.

Table 3: Fluorescence in situ hybridisation testing – costs by stage

Stage	Cost category			All categories
	Equipment	Consumables	Staff	
Slides prepared	£0.00	£0.34	£1.14	£1.48
Probe hybridisation	£4.35	£55.89	£3.05	£63.28
Fixation/counter-stain/cover slip	£0.60	£2.47	£3.14	£6.21
Analysis	£49.15	£2.95	£5.31	£57.42
Report writing	£0.00	£0.00	£4.67	£4.67
Miscellaneous	£0.43	£0.00	£0.00	£0.43
Total (before overheads)	£54.53	£61.64	£17.32	£133.49
% total cost	41%	46%	13%	-
TOTAL (including overheads calculated at 20%)	-	-	-	£160.19

Table 4: Cost of staff time for fluorescence in situ hybridisation testing

Staff member	Time (min)	Cost per patient sample at different staff costs per hour		
		Mean cost	Low cost	High cost
Medical technical officer ^a	50.56	£14.45	£12.54	£16.36
Clinical scientist ^b	6.94	£2.87	£2.47	£3.26
TOTAL	57.50	£17.32	£15.01	£19.62

^a Mean cost/hour £17.14 (Low: £14.88, High: £19.41); ^b Mean cost/hour £24.81 (Low: £21.40, High: £28.21)

3.3 Sanger sequencing

Sanger sequencing also consists of six stages (**Table 5**). In contrast with FISH testing, most of the costs relate to either consumables (52%) or staffing (40%). There is no single overriding cost driver in either of these cost groups. More detail on staff costs is provided in **Table 6**. Most of the staff costs are again MTO time.

Table 5: Sanger sequencing – costs by stage

Stage	Cost category			All categories
	Equipment	Consumables	Staff	
DNA extraction	£2.85	£4.31	£13.14	£20.30
Polymerase chain reaction (PCR) amplification	£3.12	£14.13	£3.69	£20.94
Quantification	£0.14	£8.41	£4.57	£13.12
Sequencing	£0.00	£14.69	£1.30	£15.99
Analysis	£0.44	£0.00	£9.53	£9.97
Miscellaneous	£0.23	£0.00	£0.00	£0.23
Total (before overheads)	£6.78	£41.54	£32.23	£80.55
% total cost	8%	52%	40%	-
TOTAL (including overheads calculated at 20%)	-	-	-	£96.67

Table 6: Cost of staff time for Sanger sequencing

Staff member	Time (min)	Cost per patient sample at different staff costs per hour		
		Mean cost	Low cost	High cost
Medical technical officer ^a	101.95	£29.13	£25.28	£32.98
Clinical scientist ^b	7.50	£3.10	£2.68	£3.53
TOTAL	109.45	£32.23	£27.96	£36.50

^a Mean cost/hour £17.14 (Low: £14.88, High: £19.41); ^b Mean cost/hour £24.81 (Low: £21.40, High: £28.21)

3.4 Targeted next generation sequencing

Targeted NGS (using an Illumina MiSeq Benchtop Sequencer) consists of eight stages (**Table 7**). Consumables account for most (67%) of the costs of this test. The cost of the targeted NGS kit is £10,800, which provides consumables for 96 samples (£112.50 each). The other major consumable cost is the MiSeq reagent cartridge used in the Library Pooling stage (£629.00 per cartridge). This can be used to process a batch of 24 samples, costing £26.21 per sample. The major piece of equipment used during this stage is the MiSeq Sequencer. This is estimated to cost £99,938, with an annual maintenance contract costing £11,700, and a lifespan of five years. It was estimated that currently 15-20% of MiSeq time is devoted to CLL testing; a mid-point value of 17.5% was used.

Table 7: Targeted next generation sequencing – costs by stage

Stage	Cost category			All categories
	Equipment	Consumables	Staff	
DNA extraction	£2.85	£4.31	£13.14	£20.30
Hybridisation of oligo pool	£1.27	£26.68	£0.57	£28.51
Removal of unbound oligos	£0.51	£29.39	£0.64	£30.54
Extension-ligation of bound oligos	£0.23	£6.69	£0.10	£7.01
PCR amplification	£2.97	£15.93	£0.49	£19.40
PCR clean-up	£2.06	£8.27	£0.70	£11.03
Library normalisation	£1.37	£26.52	£0.65	£28.55
Library pooling for MiSeq sequencing	£41.25	£32.87	£0.38	£74.50
Data analysis	£0.64	£11.00	£8.54	£20.18
Total (before overheads)	£53.16	£161.65	£25.23	£240.04
% total cost	22%	67%	11%	-
TOTAL (including overheads calculated at 20%)	-	-	-	£288.05

More detail on the staff costs associated with targeted NGS is provided in **Table 8**.

Table 8: Cost of staff time for targeted next generation sequencing

Staff member	Time (min)	Cost per patient sample at different staff costs per hour		
		Mean cost	Low cost	High cost
Medical technical officer ^a	58.40	£16.69	£14.48	£18.89
Clinical scientist ^b	20.67	£8.54	£7.37	£9.72
TOTAL	79.06	£25.23	£21.85	£28.61

^a Mean cost/hour £17.14 (Low: £14.88, High: £19.41); ^b Mean cost/hour £24.81 (Low: £21.40, High: £28.21)

3.5 Collation of results – Fluorescence in situ hybridisation testing and Sanger sequencing

Table 9 summarises the costs associated with collating the results from FISH testing and Sanger sequencing and preparing a joint report. Almost all of these costs (96%) are staff time (**Table 10**).

Table 9: Collation of results for fluorescence in situ hybridisation testing and Sanger sequencing - costs by stage

Stage	Cost category			All categories
	Equipment	Consumables	Staff	
Results collated and report prepared	£0.64	£0.00	£13.98	£14.62
% total cost	4%	0%	96%	-
TOTAL (including overheads calculated at 20%)	-	-	-	£17.55

Table 10: Cost of staff time for collation of results for fluorescence in situ hybridisation testing and Sanger sequencing

Staff member	Time (min)	Cost per patient sample at different staff costs per hour		
		Mean cost	Low cost	High cost
Medical technical officer ^a	20.00	£5.71	£4.96	£6.47
Clinical scientist ^b	20.00	£8.27	£7.13	£9.40
TOTAL	40.00	£13.98	£12.09	£15.87

^a Mean cost/hour £17.14 (Low: £14.88, High: £19.41); ^b Mean cost/hour £24.81 (Low: £21.40, High: £28.21)

3.6 Preparation of results –targeted NGS

Table 11 summarises the costs associated with preparing the targeted NGS results. Again, almost all of these costs are staff time (**Table 12**).

Table 11: Preparation of targeted next generation sequencing results - costs by stage

Stage	Cost category			All categories
	Equipment	Consumables	Staff	
Results collated and report prepared	£0.64	£0.00	£19.78	£20.41
% total cost	3%	0%	97%	-
TOTAL (including overheads calculated at 20%)	-	-	-	£24.50

Table 12: Cost of staff time for the preparation of targeted next generation sequencing results

Staff member	Time (min)	Cost per patient sample at different staff costs per hour		
		Mean cost	Low cost	High cost
Clinical scientist ^a	26.00	£10.75	£9.27	£12.23
Bioinformatician ^b	6.00	£2.88	£2.63	£3.13
Senior clinical researcher ^c	6.00	£6.15	£5.23	£7.06
TOTAL	38.00	£19.78	£17.14	£22.42

^a Mean cost/hour £24.81 (Low: £21.40, High: £28.21); ^b Mean cost/hour £28.80 (Low: £26.27, High: £31.34); ^c Mean cost/hour £61.46 (Low: £52.35, High: £70.57)

3.7 Total costs

Table 13 and **Table 14** summarise the total costs of testing for each pathway. **Table 15** compares costs across both pathways. The future testing pathway costs £38.15 more than the current pathway. Although the consumable costs associated with future practice are higher, these are offset by savings in equipment and staff costs. **Tables 16** and **17** summarise staff costs for both pathways. The savings in future practice are mainly driven by a large reduction in MTO time.

Table 13: Total cost of testing, current pathway

Stage	Cost category			All categories
	Equipment	Consumables	Staff	
Sample reception	£0.16	£0.02	£5.40	£5.57
FISH	£54.53	£61.64	£17.32	£133.49
Sanger sequencing	£6.78	£41.54	£32.23	£80.55
Collation of results	£0.64	£0.00	£13.98	£14.62
Total (before overheads)	£62.11	£103.21	£68.93	£234.24
% total cost	27%	44%	29%	-
TOTAL (including overheads calculated at 20%)	-	-	-	£281.08

Table 14: Total cost of testing, future pathway

Stage	Cost category			All categories
	Equipment	Consumables	Staff	
Sample reception	£0.16	£0.02	£5.40	£5.57
Targeted NGS	£53.16	£161.65	£25.23	£240.04
Preparation of results	£0.64	£0.00	£19.78	£20.41
Total (before overheads)	£53.96	£161.66	£50.40	£266.02
% total cost	20%	61%	19%	-
TOTAL (including overheads calculated at 20%)	-	-	-	£319.23

Table 15: Cost comparison between current and future testing practice

Stage	Cost category				All categories
	Equipment	Consumables	Staff	Overheads	
Current practice (FISH + Sanger)	£62.11	£103.21	£68.93	£46.85	£281.08
Future practice (Targeted NGS)	£53.96	£161.66	£50.40	£53.20	£319.23
DIFFERENCE	-£8.15	£58.46	-£18.52	£6.36	£38.15
Future cost % change	-13%	57%	-27%	14%	14%

Table 16: Total cost of staff time, current pathway

Staff member	Time (min)	Cost per patient sample at different staff costs per hour		
		Mean cost	Low cost	High cost
Medical technical officer ^a	189.64	£54.19	£47.03	£61.35
Clinical scientist ^b	34.44	£14.24	£12.28	£16.19
Principle clinical scientist ^c	1.00	£0.50	£0.45	£0.55
TOTAL	225.08	£68.93	£59.76	£78.09

^a Mean cost/hour £17.14 (Low: £14.88, High: £19.41); ^b Mean cost/hour £24.81 (Low: £21.40, High: £28.21); ^c Mean cost/hour £30.03 (Low: £27.30, High: £32.76)

Table 17: Total cost of staff time, future pathway

Staff member	Time (min)	Cost per patient sample at different staff costs per hour		
		Mean cost	Low cost	High cost
Medical technical officer ^a	75.53	£21.58	£18.73	£24.43
Clinical scientist ^b	46.67	£19.29	£16.65	£21.94
Principle clinical scientist ^c	1.00	£0.50	£0.45	£0.55
Bioinformatician ^d	6.00	£2.88	£2.63	£3.13
Senior clinical researcher ^e	6.00	£6.15	£5.23	£7.06
TOTAL	135.20	£50.40	£43.69	£57.11

^a Mean cost/hour £17.14 (Low: £14.88, High: £19.41); ^b Mean cost/hour £24.81 (Low: £21.40, High: £28.21); ^c Mean cost/hour £30.03 (Low: £27.30, High: £32.76); ^d Mean cost/hour £28.80 (Low: £26.27, High: £31.34); ^e Mean cost/hour £61.46 (Low: £52.35, High: £70.57)

3.8 Sensitivity analysis

Table 18 describes the results of the sensitivity analysis. Sample throughput has a notable impact on the costing results. An increase from 150 samples a year to 250 increases the cost difference between current and future practice to £70.07, an increase of 84%. This is because the impact of an increase in throughput on equipment costs is relatively higher in current practice, where equipment costs account for more of the total test cost.

A number of the parameters related to staffing are also important. If staff output is higher (more weeks worked, or more hours worked per week), this favours current practice, which requires more staff time. Using clinical scientists (higher grade staff) to perform the tasks usually completed by MTOs also impacts on the results, making genomic testing the cheaper option by £20.66.

Finally, the results are very sensitive to the costs of the equipment and consumables used in future practice, and the assumptions that are made regarding the use of these items. A 33% reduction in the cost of the kit used for targeted NGS leads to a 119% reduction in the cost difference between the two tests, with genomic testing becoming £7.15 cheaper. The cost of the MiSeq system itself is important, but not as important as the proportion of MiSeq time that is used for CLL testing. This is 17.5% in the base case. If this increases to 30%, genomic testing becomes £71.98 more expensive (an increase of 89%), as more of the MiSeq costs are attributed to CLL testing.

Table 18: Sensitivity analysis results

Parameter (base case value)	Variation	Cost (% increase/decrease compared to base case costs)		
		Current practice	Future practice	Difference
BASE CASE RESULTS		£281.08	£319.23	£38.15
PARAMETERS AFFECTING BOTH CURRENT AND FUTURE TESTING PRACTICE				
CLL samples per year (150)	50	£570.23 (103%)	£448.73 (41%)	£-121.50 (-419%)
	100	£353.37 (26%)	£351.60 (10%)	£-1.77 (-105%)
	200	£244.94 (-13%)	£303.04 (-5%)	£58.10 (52%)
	250	£223.26 (-21%)	£293.33 (-8%)	£70.07 (84%)
National Insurance / Superannuation multiplier (20%)	0%	£267.30 (-5%)	£309.15 (-3%)	£41.85 (10%)
	10%	£274.19 (-2%)	£314.19 (-2%)	£40.00 (5%)
	30%	£287.98 (2%)	£324.27 (2%)	£36.29 (-5%)
	40%	£294.87 (5%)	£329.31 (3%)	£34.44 (-10%)
Weeks worked per year (46)	42	£288.96 (3%)	£324.99 (2%)	£36.03 (-6%)
	44	£284.84 (1%)	£321.98 (1%)	£37.14 (-3%)
	48	£277.64 (-1%)	£316.71 (-1%)	£39.07 (2%)
	50	£274.47 (-2%)	£314.39 (-2%)	£39.92 (5%)
Hours worked per week (37.5)	30	£301.76 (7%)	£334.35 (5%)	£32.59 (-15%)
	35	£286.99 (2%)	£323.55 (1%)	£36.56 (-4%)
	40	£275.91 (-2%)	£315.45 (-1%)	£39.53 (4%)
	45	£267.30 (-5%)	£309.15 (-3%)	£41.85 (10%)
Overheads (20%)	0%	£234.24 (-17%)	£266.02 (-17%)	£31.79 (-17%)
	10%	£257.66 (-8%)	£292.63 (-8%)	£34.97 (-8%)
	30%	£304.51 (8%)	£345.83 (8%)	£41.32 (8%)
Discount rate (3.5%)	2%	£278.46 (-1%)	£317.29 (-1%)	£38.83 (2%)
	6%	£285.60 (2%)	£322.58 (1%)	£36.98 (-3%)
	10%	£293.07 (4%)	£328.16 (3%)	£35.09 (-8%)
Assumed equipment lifespan when information unavailable (five years)	10	£276.44 (-2%)	£315.54 (-1%)	£39.10 (3%)
	15	£274.90 (-2%)	£314.32 (-2%)	£39.42 (3%)
Clinical scientists used instead of MTOs	-	£310.15 (10%)	£330.81 (4%)	£20.66 (-46%)
MTOs used instead of clinical scientists	-	£275.81 (-2%)	£312.08 (-2%)	£36.27 (-5%)
Consultant clinical scientists used instead of clinical scientists	-	£299.34 (6%)	£343.97 (8%)	£44.63 (17%)
Clinical scientists used instead of consultant clinical scientists	-	£281.08 (0%)	£319.23 (0%)	£38.15 (0%)
Low point of MTO salary range used instead of mid-point	-	£272.49 (-3%)	£315.81 (-1%)	£43.32 (14%)
High point of MTO salary range instead of mid-point	-	£289.68 (3%)	£322.65 (1%)	£32.98 (-14%)
Low point of clinical scientist salary range instead of mid-point	-	£278.74 (-1%)	£316.05 (-1%)	£37.31 (-2%)
High point of clinical scientist salary range instead of mid-point	-	£283.43 (1%)	£322.41 (1%)	£38.98 (2%)
PARAMETERS AFFECTING CURRENT PRACTICE ONLY				
FISH testing batch size (3)	2	£284.98 (1%)	£319.23 (0%)	£34.25 (-10%)
	4	£279.13 (-1%)	£319.23 (0%)	£40.10 (5%)
Sanger sequencing batch size (4)	2	£303.99 (8%)	£319.23 (0%)	£15.24 (-60%)
	6	£273.45 (-3%)	£319.23 (0%)	£45.78 (20%)
Acetone cost (£130)	£100	£266.57 (-5%)	£319.23 (0%)	£52.66 (38%)
	£160	£295.60 (5%)	£319.23 (0%)	£23.63 (-38%)

Parameter (base case value)	Variation	Cost (% increase/decrease compared to base case costs)		
		Current practice	Future practice	Difference
BASE CASE RESULTS		£281.08	£319.23	£38.15
Fluorescent microscope cost (£57,500)	£45,000	£270.45 (-4%)	£319.23 (0%)	£48.77 (28%)
	£70,000	£291.71 (4%)	£319.23 (0%)	£27.52 (-28%)
Sequencing cost (£75)	£50	£276.03 (-2%)	£319.23 (0%)	£43.20 (13%)
	£100	£286.13 (2%)	£319.23 (0%)	£33.10 (-13%)
Average number of samples sequenced per plate (6)	3	£296.23 (5%)	£319.23 (0%)	£23.00 (-40%)
	8	£277.30 (-1%)	£319.23 (0%)	£41.93 (10%)
PARAMETERS AFFECTING FUTURE PRACTICE ONLY				
Targeted NGS batch size (24)	16	£281.08 (0%)	£339.27 (6%)	£58.18 (53%)
	32	£281.08 (0%)	£309.21 (-3%)	£28.13 (-26%)
Kit cost for targeted NGS (£9,000)	£6,000	£281.08 (0%)	£273.94 (-14%)	£-7.15 (-119%)
	£12,000	£281.08 (0%)	£364.52 (14%)	£83.44 (119%)
MiSeq cartridge cost (£524)	£400	£281.08 (0%)	£311.78 (-2%)	£30.70 (-20%)
	£700	£281.08 (0%)	£329.78 (3%)	£48.70 (28%)
MiSeq system cost (£83,282)	£75,000	£281.08 (0%)	£316.15 (-1%)	£35.06 (-8%)
	£85,000	£281.08 (0%)	£319.87 (0%)	£38.78 (2%)
	£115,000	£281.08 (0%)	£331.03 (4%)	£49.95 (31%)
	£125,000	£281.08 (0%)	£334.75 (5%)	£53.67 (41%)
% MiSeq time used for CLL testing (17.5%)	5%	£281.08 (0%)	£285.40 (-11%)	£4.31 (-89%)
	10%	£281.08 (0%)	£298.93 (-6%)	£17.85 (-53%)
	15%	£281.08 (0%)	£312.46 (-2%)	£31.38 (-18%)
	25%	£281.08 (0%)	£339.53 (6%)	£58.44 (53%)
	30%	£281.08 (0%)	£353.06 (11%)	£71.98 (89%)
Data storage time period (30 years)	10	£281.08 (0%)	£312.03 (-2%)	£30.95 (-19%)
	20	£281.08 (0%)	£315.63 (-1%)	£34.55 (-9%)
	40	£281.08 (0%)	£322.83 (1%)	£41.75 (9%)

CLL = chronic lymphocytic leukaemia; FISH = fluorescent in situ hybridisation; MTO = medical technical officer; NGS = next generation sequencing.

Appendix 18: Complication rates

Treatment	Line of therapy / patient subgroup	Average number of complications experienced per 28 day cycle of treatment	Notes	Sources
FCR	First-line	0.3489	733 grade 3/4 events experienced across 2101 treatment courses	[230]
	Second-line	0.6978	No information was available in the literature on this parameter. Clinical experience at OUH NHS Trust suggests that the complication rate doubles with each line of treatment. The first-line complication rate was therefore doubled	Ass
	First-line – FCR responders	0.1745	No information was available in the literature on this parameter. It was therefore assumed that the complication rate would halve due to improved targeting of treatment	Ass
BR	First-line	0.3626	No information was available in the literature on this parameter. Clinical experience at OUH NHS Trust suggests that the complication rate doubles with each line of treatment. The second-line complication rate was therefore halved	Ass
	Second-line	0.7252	256 grade 3/4 events experienced across 353 treatment courses	[328]
	Second-line – FCR responders	0.3626	No information was available in the literature on this parameter. It was therefore assumed that the complication rate would halve due to improved targeting of treatment	Ass
Ofa	First-line	0.0922	No information was available in the literature on this parameter. Clinical experience at OUH NHS Trust suggests that the complication rate doubles with each line of treatment. The second-line complication rate was therefore halved	Ass
	Second- or third-line	0.1843	Byrd <i>et al.</i> [219] report 82 grade 3/4 events experienced across approximately 1100 treatment courses. A second study [197] reports 85 grade 3/4 events experienced across approximately 289 treatment courses. An unweighted average was taken of these two complication rates	[197,219]

Treatment	Line of therapy / patient subgroup	Average number of complications experienced per 28 day cycle of treatment	Notes	Sources
IB	Third-line	0.0581	106 grade 3/4 events experienced across approximately 1823 treatment courses	[219]
	Third-line – FCR responders	0.0291	No information was available in the literature on this parameter. It was therefore assumed that the complication rate would halve due to improved targeting of treatment	Ass
	First-line – FCR non-responders	0.0291	No information was available in the literature on this parameter. It was therefore assumed that the complication rate would be the same as that in FCR responders receiving third-line IB treatment	Ass
BMT	-	0.0195	This complication rate was estimated based on clinical experience at OUH NHS Trust, which suggested that 80% of patients experience one complication whilst in remission (an average of 51 28-day cycles)	Ass

Ass = assumption; BMT = bone marrow transplant; BR = bendamustine and rituximab; FCR = fludarabine, cyclophosphamide and rituximab; IB = ibrutinib; Ofa = ofatumumab; OUH = Oxford University Hospitals.

Appendix 19: Costs associated with treating chemotherapy complications

Chemotherapy complication	Unit cost	Notes	Source
Anaemia	£2,254	Currency codes SA03G/H: haemolytic anaemia. Average used of the costs for clinical haematology, weighted by activity	[232]
CMV reactivation	£1,042	Ganciclovir treatment. Assumed an average of 35 infusions per course of treatment	[234]
Dyspnoea	£3,409	Currency codes DZ21A/M/Q/R/S/T/U: chronic obstructive pulmonary disease or bronchitis. Average used of the costs for clinical haematology, weighted by activity	[232,333]
Febrile neutropenia	£4,009	In the absence of specific cost data, this cost was assumed to be equal to that for infections (see below)	[232]
Fever	£1,608	Currency code WA04Z: acute febrile illness with length of stay 4 days or less – clinical haematology	[232]
Infection (including pneumonia)	£4,009	Currency codes DZ11-J (lobar, atypical or viral pneumonia) and WA03A-C (septicaemia). Average used of the costs for clinical haematology, weighted by activity	[232]
Thrombocytopenia	£3,106	Currency code SA12G-K: thrombocytopenia. Average used of the costs for clinical haematology, weighted by activity	[232]
Unweighted average cost	£2,777		

CMV = cytomegalovirus.

Appendix 20: Costs associated with treating bone marrow transplant complications

BMT complication	Unit cost	Notes	Source
Infection (including pneumonia)	£4,009	Currency codes DZ11-J (lobar, atypical or viral pneumonia) and WA03A-C (septicaemia). Average used of the costs for clinical haematology, weighted by activity	[232]
Chronic graft versus host disease	£477	Unweighted average cost calculated for three treatment options: (1) Steroids – 2mg/kg of methylprednisolone daily for 28 days; (2) PUVA light therapy; (3) Photopheresis (Currency code for (2) and (3): JC47A)	[232]
CMV reactivation	£1,042	Ganciclovir treatment. Assumed an average of 35 infusions per course of treatment	[234]
Unweighted average cost	£1,843		

BMT = bone marrow transplant; CMV = cytomegalovirus.

Appendix 21: Average length of chemotherapy treatment

Disease state	Average number of 28-day cycles spent in state	Notes	Source
Undergoing FCR treatment	5.2	Mean number of treatment courses	[230]
Undergoing BR treatment	4.5	Mean number of treatment courses	[328]
Undergoing Ofa treatment	5.8	Median treatment exposure was reported to be 5.3 months	[219]
Undergoing IB treatment	9.3	Median treatment exposure was reported to be 8.6 months	[219]
In remission following FCR treatment	56.3	Median PFS was reported to be 51.8 months	[230]
In remission following BMT	51.3	The relapse incidence at 4 years following BMT was reported to be 0.4. This was converted into a 28-day transition probability of 0.0097 (see p17, p27, p51, p60 in Appendix 14). This transition probability suggests that it would take 102.6 28-day cycles for all patients to exit remission following BMT. It was assumed that half of all patients would have exited this state exactly halfway through this period i.e. after 51.3 cycles	[330]

BMT = bone marrow transplant; BR = bendamustine and rituximab; BSC = best supportive care; FCR = fludarabine, cyclophosphamide and rituximab; IB = ibrutinib; Ofa = ofatumumab; PFS = progression-free survival.

Appendix 22: Structured literature search conducted to identify utility values for chronic lymphocytic leukaemia patients in different disease states

Background and methods

The aim of this literature review was to identify references that report utility values for the various disease states that could be experienced by chronic lymphocytic leukaemia (CLL) patients. These utility values were then used as parameters in the model that was constructed to evaluate the cost-effectiveness and cost-utility of using genomic testing to guide treatment decisions in CLL. An earlier study conducted a literature review with a similar objective, identifying a number of utility values with potential relevance for this model [238]. This earlier literature review considered all references published on or before 2nd June 2010. To ensure that this model made use of all relevant utility information in this context, the same search syntax was applied from 3rd June 2010 until 13th June 2014.

The search strategies were described in detail in the previous study. In brief, the strategies combined multiple keyword variations using the logical operator 'OR' to identify studies reporting utility values for health states or toxicities associated with CLL and CLL treatment, or cost-effectiveness analyses, health technology assessment reports and review articles potentially reporting relevant utilities. Only English language studies were considered. Literature searches were undertaken in Embase, Medline and Medline In-Process, Econlit, and The UK National Health Service (NHS) Economic Evaluation Database. The reference lists of included articles were manually reviewed to check for additional studies that might also meet the inclusion criteria. In addition, a manual search of the NICE website was undertaken to identify relevant single technology appraisals.

The search syntax used in each database is listed below.

Medline and Medline in-process

Leukemia, Lymphocytic, Chronic, B-Cell/
cll.tw.
(chronic\$ and lymph\$ and leuk?em\$).tw.
(chronic\$ and lymph\$ and leuc?em\$).tw.
or/1-4
quality - adjusted life years/
quality adjusted life.tw.
quality of well being.tw.
quality of wellbeing.tw.
qwb.tw.
index of wellbeing.tw.
index of well being.tw.
(health related quality adj2 life\$).tw.
classification of illness state\$.tw.
(standard gamble or sg).tw.
(timetradeoff or time tradeoff or time trade off or tto).tw.
(health adj3 (indicator? or status)).tw.
(qaly? or qale? or hrqol? or hrql).tw.

(euroqol or euroqol or euro qol or eq 5d or eq5d).tw.
 (hui2 or hui 2 or hui3 or hui 3 or hui).tw.
 (sf 36 or sf36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or short form thirty six or short form thirtysix or short form thirty six).tw.
 (sf 6d or sf6d or short form 6d or shortform 6d or sf six\$ or shortform six\$ or short form six\$).tw.
 (fact adj1 leu).tw.
 functional assessment of cancer therapy.tw.
 (qlq c30 or qlq c 30).tw.
 qlq leu.tw.
 exp Health Status Indicators/
 (cost\$ adj1 (effective\$ or benefit\$)).tw.
 Pain Measurement/
 visual analogue scale\$.tw.
 vas.tw.
 "Activities of Daily Living"/
 utilit\$.tw.
 or/6 - 33
 limit 35 to english language
 limit 36 to yr="2010 -Current"

Embase 1996 to 2014 Week 23

chronic lymphatic leukemia/
 cll.tw
 (chronic\$ and lymph\$ and leuk?em\$).tw.
 (chronic\$ and lymph\$ and leuc?em\$).tw.
 or/1-4
 quality adjusted life year/
 quality adjusted life.tw.
 quality of well being.tw.
 quality of wellbeing.tw.
 qwb.tw.
 index of wellbeing.tw.
 index of well being.tw.
 (health related quality adj2 life\$).tw.
 classification of illness state\$.tw.
 (standard gamble or sg).tw.
 wellbeing/
 time trade off/
 (timetradeoff or time tradeoff or time trade off or tto).tw.
 (health adj3 (indicator? or status)).tw.
 (qaly? or qale? or hrqol? or hrql).tw.
 (euroqol or euroqol or euro qol or eq 5d or eq5d).tw.
 (hui2 or hui 2 or hui3 or hui 3 or hui).tw.
 (sf 36 or sf36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or short form thirty six or short form thirtysix or short form thirty six).tw.
 (sf 6d or sf6d or short form 6d or shortform 6d or sf six\$ or shortform six\$ or short form six\$).tw.
 (fact adj1 leu).tw.
 functional assessment of cancer therapy.tw.
 (qlq c30 or qlq c 30).tw.

qlq leu.tw.
 European quality of life questionnaire 5D/
 european quality of life 5 dimension/
 short form 36/
 exp health status/
 (cost\$ adj1 (effective\$ or benefit\$)).tw.
 visual analog scale/
 visual analogue scale\$.tw.
 vas.tw.
 daily life activity/
 utilit\$.tw.
 or/6-38
 5 and 39
 limit 40 to english language
 limit 41 to yr="2010 -Current"

EconLIT

As the EconLIT database is now provided by ProQuest rather than Ovid, the search strategy that was used previously was updated to fit the new system, and restricted to 2010-2014:

leukaemia OR leukaemias OR leukemia OR leukemias OR leucaemia OR leucaemias OR leucemia OR leucemias OR cll

NHS EED

MeSH DESCRIPTOR Leukemia, Lymphocytic, Chronic, B-Cell EXPLODE 1 2 3
 Cll

Chronic* AND lymph* AND leukem*
 Chronic* AND lymph* AND leukaem*
 Chronic* AND lymph* AND leucaem*
 Chronic* AND lymph* AND leucem*

Results

The search results are described in the following table:

Database	Date	References identified	After removing duplicates	After reviewing titles	After reviewing abstracts	Articles evaluated, after discarding irrelevant and unavailable articles, and adding 'other' references
Embase	13/06/14	463	452	159	26	3
Medline	13/06/14	124	29	10	3	0
Econlit	13/06/14	9	8	0	0	0
NHS EED	13/06/14	14	1	0	0	0
Other	-	-	-	-	-	4
Total		610	490	169	29	7
References not retained			120	321	140	24

NHS EED = NHS economic evaluation database.

A large number of those studies initially identified were not retained in the final count because these studies were all economic evaluations that used similar data sources for the utility parameters. In addition, a number of the 'Other' references included in the final count were not identified in the four specified databases because they were submitted as evidence to the National Institute for Health and Care Excellence instead of being published in a peer-reviewed journal.

After removing duplicates, reviewing titles, abstracts and full articles, and adding references from other sources, seven studies were taken forward to the next step in the modelling process, in which relevant utility data was extracted [238,239,319,337-340]. This utility data was then combined with that identified in the earlier literature review, and the most appropriate utility estimates were selected for use in the modelling process.

Appendix 23: Parameter variations considered in univariate sensitivity analysis

Parameter variation	Value			Justification for low/high values
	Base	Low	High	
Test costs				
Cost of genetic testing (£)	281.08	140.54	421.62	Assumed 50% variation
Cost of genomic testing (£)	319.23	159.62	478.85	Assumed 50% variation
Cost of a FBC (£)	3.01	1.68	3.97	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost of a U&E test (£)	1.25	0.82	1.43	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost of a LFT (£)	1.25	0.82	1.43	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost of a CT scan (£)	106.32	89.71	143.10	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost of an IgVH test (£)	38.74	16.50	51.57	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost of a simple lung function test (£)	40.61	29.69	37.34	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost of a bone marrow aspirate (£)	482.48	313.39	590.59	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost of a thyroid function test (£)	1.25	0.82	1.43	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost of an immunoglobulin test (£)	3.01	1.68	3.97	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost of chemotherapy treatment				
FCR treatment – first cycle (£)	2,554.74	1,277.37	3,832.12	Assumed 50% variation
FCR treatment – subsequent cycles (£)	2,876.02	1,438.01	4,314.03	Assumed 50% variation
BR treatment – first cycle (£)	2,854.63	1,427.32	4,281.95	Assumed 50% variation
BR treatment – subsequent cycles (£)	3,165.66	1,582.83	4,748.49	Assumed 50% variation
Ofatumumab treatment – first cycle (£)	11,826.62	5,913.31	17,739.94	Assumed 50% variation
Ofatumumab treatment – second cycle (£)	14,892.64	7,446.32	22,338.96	Assumed 50% variation
Ofatumumab treatment – subsequent cycles (£)	3,941.89	1,970.95	5,912.84	Assumed 50% variation
Ibrutinib treatment (£)	2,472.10	1,236.05	3,708.15	Assumed 50% variation
Other drug costs				
Cost per 960mg dose of co-trimoxazole (£)	0.19	0.10	0.29	Assumed 50% variation

Parameter variation	Value			Justification for low/high values
	Base	Low	High	
Cost per 50mg dose of flucanazole (£)	0.03	0.01	0.04	Assumed 50% variation
Cost per 200mg dose of aciclovir (£)	0.03	0.02	0.05	Assumed 50% variation
Cost per 300mcg injection of GCSF (£)	55.78	27.89	83.66	Assumed 50% variation
Cost per 100mg injection of hydrocortisone (£)	1.08	0.54	1.62	Assumed 50% variation
Cost per 30mg dose of chlorpheniramine (£)	0.06	0.03	0.09	Assumed 50% variation
Cost per 1g dose of paracetamol (£)	0.01	0.01	0.02	Assumed 50% variation
Other costs				
Cost for IVIG treatment (£)	1,201.23	600.62	1,801.85	Assumed 50% variation
Cost for a haematology appointment (consultant) (£)	142.71	97.36	171.55	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost for a haematology appointment (non-consultant) (£)	90.72	53.1	111.75	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost for a telephone consultation with a haematology nurse (£)	60.77	39.39	62.03	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost for an immunology appointment (consultant) (£)	195.01	85.24	322.01	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost for an inpatient stay (£)	2,577.16	1,708.39	3,110.34	Informed by the lower and upper quartiles in the NHS Reference Costs database [232]
Cost to treat a complication arising from chemotherapy (£)	2,776.70	1,388.35	4,165.04	Assumed 50% variation
Cost per BMT (£)	43,723.96	21,861.98	65,585.93	Assumed 50% variation
Cost to treat a complication arising from BMT (£)	1,842.66	921.33	2,763.99	Assumed 50% variation
Cost per donor lymphocyte infusion (£)	381.88	190.94	572.82	Assumed 50% variation
Transition probabilities				
Comparator A/B: proportion of patients without <i>TP53</i> mutations	0.9323	0.8500	0.9500	Assumption
Comparator A/B: proportion of patients receiving FCR as first-line treatment	0.8000	0.7500	0.8500	Assumption
Comparator A/B: proportion of patients receiving FCR as second-line treatment	0.2000	0.1500	0.2500	Assumption
Proportion of patients who move from second-line FCR/BR straight to BSC: no <i>TP53</i> mutation/FCR responder	0.0500	0.0250	0.0750	Assumption
Proportion of patients who move from first-line Ofa treatment straight to BSC: <i>TP53</i> mutation/FCR non-responder	0.0500	0.0250	0.0750	Assumption

Parameter variation	Value			Justification for low/high values
	Base	Low	High	
28-day TP of dying from CLL during first-line FCR treatment: no <i>TP53</i> mutation	0.0027	0.0014	0.0041	Assumed 50% variation
28-day TP of no response when undergoing first-line FCR treatment: no <i>TP53</i> mutation	0.0016	0.0008	0.0024	Assumed 50% variation
28-day TP of patients moving straight to refractory treatment as high risk genetic mutations emerge following first-line FCR treatment: no <i>TP53</i> mutation	0.0006	0.0003	0.0009	Assumed 50% variation
28-day TP of dying from CLL during first-line BR treatment: no <i>TP53</i> mutation	0.0027	0.0014	0.0041	Assumed 50% variation
28-day TP of no response when undergoing first-line BR treatment: no <i>TP53</i> mutation	0.0016	0.0008	0.0024	Assumed 50% variation
28-day TP of patients moving straight to refractory treatment as high risk genetic mutations emerge following first-line BR treatment: no <i>TP53</i> mutation	0.0006	0.0003	0.0009	Assumed 50% variation
28-day TP of dying from CLL during second-line FCR treatment: no <i>TP53</i> mutation	0.0033	0.0016	0.0049	Assumed 50% variation
28-day TP of moving to refractory treatment following no response to first-line FCR treatment: no <i>TP53</i> mutation	0.0049	0.0025	0.0074	Assumed 50% variation
28-day TP of moving to BSC following no response to first-line FCR treatment: no <i>TP53</i> mutation	0.0003	0.0001	0.0004	Assumed 50% variation
28-day TP of dying from CLL during second-line BR treatment: no <i>TP53</i> mutation	0.0178	0.0089	0.0266	Assumed 50% variation
28-day TP of no response when undergoing second-line BR treatment: no <i>TP53</i> mutation	0.0049	0.0025	0.0074	Assumed 50% variation
28-day TP of moving to BSC following second-line BR treatment: no <i>TP53</i> mutation	0.0003	0.0001	0.0004	Assumed 50% variation
28-day TP of dying from CLL during refractory Ofa treatment: no <i>TP53</i> mutation	0.0288	0.0160	0.0416	Low and high values based on the estimates from Byrd <i>et al.</i> [219] and Wierda <i>et al.</i> [329] that informed the base case TP
28-day TP of moving to BSC following refractory Ofa treatment: no <i>TP53</i> mutation	0.0201	0.0101	0.0302	Assumed 50% variation
28-day TP of undergoing BMT following refractory Ofa treatment: no <i>TP53</i> mutation	0.0174	0.0087	0.0261	Assumed 50% variation

Parameter variation	Value			Justification for low/high values
	Base	Low	High	
28-day TP of dying from CLL during BMT remission: no <i>TP53</i> mutation	0.0082	0.0058	0.0121	Informed by the 95% CIs reported by Dreger <i>et al.</i> [330]
28-day TP of relapsing during BMT remission: no <i>TP53</i> mutation	0.0097	0.0058	0.0148	Informed by the 95% CIs reported by Dreger <i>et al.</i> [330]
28-day TP of moving to second-line treatment when in remission following first-line FCR/BR treatment: no <i>TP53</i> mutation	0.0084	0.0042	0.0126	Assumed 50% variation
28-day TP of moving to refractory treatment when in remission following second-line FCR/BR treatment: no <i>TP53</i> mutation	0.0185	0.0092	0.0277	Assumed 50% variation
28-day TP of moving to BSC when in remission following refractory Ofa treatment: no <i>TP53</i> mutation	0.0737	0.0369	0.1106	Assumed 50% variation
28-day TP of dying from CLL when receiving BSC: no <i>TP53</i> mutation	0.0753	0.0377	0.1130	Assumed 50% variation
28-day TP of dying from CLL during first-line Ofa treatment: <i>TP53</i> mutation	0.0288	0.0144	0.0432	Assumed 50% variation
28-day TP of no response when undergoing first-line Ofa treatment: <i>TP53</i> mutation	0.0440	0.0220	0.0660	Assumed 50% variation
28-day TP of undergoing BMT following refractory Ofa treatment: <i>TP53</i> mutation	0.0174	0.0087	0.0261	Assumed 50% variation
28-day TP of moving to BSC following first-line Ofa treatment: <i>TP53</i> mutation	0.0023	0.0012	0.0035	Assumed 50% variation
28-day TP of dying from CLL during BMT remission: <i>TP53</i> mutation	0.0082	0.0058	0.0121	Informed by the 95% CIs reported by Dreger <i>et al.</i> [330]
28-day TP of relapsing during BMT remission: <i>TP53</i> mutation	0.0097	0.0058	0.0148	Informed by the 95% CIs reported by Dreger <i>et al.</i> [330]
28-day TP of dying from CLL during refractory Ofa treatment: <i>TP53</i> mutation	0.0288	0.0160	0.0416	Low and high values based on the estimates from Byrd <i>et al.</i> [219] and Wierda <i>et al.</i> [329] that informed the base case TP
28-day TP of moving to BSC following refractory Ofa treatment: <i>TP53</i> mutation	0.0463	0.0231	0.0694	Assumed 50% variation
28-day TP of moving to refractory Ofa treatment when in remission following first-line Ofa treatment: <i>TP53</i> mutation	0.0737	0.0369	0.1106	Assumed 50% variation
28-day TP of moving to BSC when in remission following refractory Ofa treatment: <i>TP53</i> mutation	0.0737	0.0369	0.1106	Assumed 50% variation
28-day TP of dying from CLL when receiving BSC: <i>TP53</i> mutation	0.0753	0.0377	0.1130	Assumed 50% variation

Parameter variation	Value			Justification for low/high values
	Base	Low	High	
28-day TP of dying from CLL during refractory IB treatment: no <i>TP53</i> mutation	0.0032	0.0016	0.0048	Assumed 50% variation
28-day TP of moving to BSC following refractory IB treatment: no <i>TP53</i> mutation	0.0043	0.0021	0.0064	Assumed 50% variation
28-day TP of dying from CLL during refractory IB treatment: <i>TP53</i> mutation	0.0131	0.0066	0.0197	Assumed 50% variation
28-day TP of moving to BSC when in remission following refractory IB treatment: <i>TP53</i> mutation	0.0223	0.0111	0.0334	Assumed 50% variation
28-day TP of dying from CLL during first-line FCR treatment: any <i>TP53</i> mutation status	0.0035	0.0018	0.0053	Assumed 50% variation
28-day TP of no response when undergoing first-line FCR treatment: any <i>TP53</i> mutation status	0.0026	0.0013	0.0038	Assumed 50% variation
28-day TP of dying from CLL during first-line BR treatment: any <i>TP53</i> mutation status	0.0035	0.0018	0.0053	Assumed 50% variation
28-day TP of no response when undergoing first-line BR treatment: any <i>TP53</i> mutation status	0.0026	0.0013	0.0038	Assumed 50% variation
28-day TP of dying from CLL during second-line FCR treatment: any <i>TP53</i> mutation status	0.0041	0.0021	0.0062	Assumed 50% variation
28-day TP of moving to refractory treatment following no response to first-line FCR treatment: any <i>TP53</i> mutation status	0.0055	0.0027	0.0082	Assumed 50% variation
28-day TP of moving to BSC following no response to first-line FCR treatment: any <i>TP53</i> mutation status	0.0003	0.0001	0.0004	Assumed 50% variation
28-day TP of dying from CLL during second-line BR treatment: any <i>TP53</i> mutation status	0.0186	0.0093	0.0280	Assumed 50% variation
28-day TP of no response when undergoing second-line BR treatment: any <i>TP53</i> mutation status	0.0055	0.0027	0.0082	Assumed 50% variation
28-day TP of moving to BSC following second-line BR treatment: any <i>TP53</i> mutation status	0.0003	0.0001	0.0004	Assumed 50% variation
28-day TP of dying from CLL during refractory Ofa treatment: any <i>TP53</i> mutation status	0.0288	0.0160	0.0416	Low and high values based on the estimates from Byrd <i>et al.</i> [219] and Wierda <i>et al.</i> [329] that informed the base case TP
28-day TP of moving to BSC following refractory Ofa treatment: any <i>TP53</i> mutation status	0.0250	0.0125	0.0375	Assumed 50% variation

Parameter variation	Value			Justification for low/high values
	Base	Low	High	
28-day TP of undergoing BMT following refractory Ofa treatment: any <i>TP53</i> mutation status	0.0174	0.0087	0.0261	Assumed 50% variation
28-day TP of dying from CLL during BMT remission: any <i>TP53</i> mutation status	0.0082	0.0058	0.0121	Informed by the 95% CIs reported by Dreger <i>et al.</i> [330]
28-day TP of relapsing during BMT remission: any <i>TP53</i> mutation status	0.0097	0.0058	0.0148	Informed by the 95% CIs reported by Dreger <i>et al.</i> [330]
28-day TP of moving to second-line treatment when in remission following first-line FCR/BR treatment: any <i>TP53</i> mutation status	0.0093	0.0046	0.0139	Assumed 50% variation
28-day TP of moving to refractory treatment when in remission following second-line FCR/BR treatment: any <i>TP53</i> mutation status	0.0218	0.0109	0.0327	Assumed 50% variation
28-day TP of moving to BSC when in remission following refractory Ofa treatment: any <i>TP53</i> mutation status	0.0737	0.0369	0.1106	Assumed 50% variation
28-day TP of dying from CLL when receiving BSC: any <i>TP53</i> mutation status	0.0753	0.0377	0.1130	Assumed 50% variation
Intervention 1/2: proportion of patients who are FCR responders	0.8333	0.4167	1.2500	Assumed 50% variation
Intervention 1/2: proportion of patients who are correctly identified as FCR responders	0.7455	0.3727	1.1182	Assumed 50% variation
28-day TP of dying from CLL during first-line FCR treatment: FCR responder, correctly identified	0.0022	0.0011	0.0032	Assumed 50% variation
28-day TP of no response when undergoing first-line FCR treatment: FCR responder, correctly identified	0.0018	0.0009	0.0026	Assumed 50% variation
28-day TP of patients moving straight to refractory treatment as high risk genetic mutations emerge following first-line FCR treatment: FCR responder, correctly identified	0.0006	0.0003	0.0009	Assumed 50% variation
28-day TP of dying from CLL during first-line FCR treatment: FCR responder, incorrectly identified	0.0244	0.0122	0.0366	Assumed 50% variation
28-day TP of no response when undergoing first-line FCR treatment: FCR responder, incorrectly identified	0.0097	0.0049	0.0146	Assumed 50% variation
28-day TP of patients moving straight to refractory treatment as high risk genetic mutations emerge following first-line FCR treatment: FCR responder, incorrectly identified	0.0006	0.0003	0.0009	Assumed 50% variation
28-day TP of dying from CLL during second-line BR treatment: FCR responder	0.0128	0.0064	0.0192	Assumed 50% variation

Parameter variation	Value			Justification for low/high values
	Base	Low	High	
28-day TP of no response when undergoing second-line BR treatment: FCR responder	0.0028	0.0014	0.0042	Assumed 50% variation
28-day TP of moving to BSC following second-line BR treatment: FCR responder	0.0012	0.0006	0.0018	Assumed 50% variation
28-day TP of dying from CLL during refractory IB treatment: FCR responder	0.0032	0.0016	0.0048	Assumed 50% variation
28-day TP of moving to second-line treatment when in remission following first-line FCR treatment: FCR responder, correctly identified	0.0070	0.0035	0.0105	Assumed 50% variation
28-day TP of moving to second-line treatment when in remission following first-line FCR treatment: FCR responder, incorrectly identified	0.0361	0.0180	0.0541	Assumed 50% variation
28-day TP of moving to refractory treatment when in remission following second-line BR treatment: FCR responder	0.0127	0.0064	0.0191	Assumed 50% variation
28-day TP of moving to BSC following refractory IB treatment: FCR responder	0.0043	0.0021	0.0064	Assumed 50% variation
28-day TP of dying from CLL when receiving BSC: FCR responder	0.0753	0.0377	0.1130	Assumed 50% variation
28-day TP of dying from CLL during first-line IB treatment: FCR non responder	0.0025	0.0012	0.0037	Assumed 50% variation
28-day TP of moving to BSC following first-line IB treatment: FCR non-responder	0.0041	0.0020	0.0061	Assumed 50% variation
28-day TP of dying from CLL when receiving BSC: FCR non-responder	0.0753	0.0377	0.1130	Assumed 50% variation
Multiplier for mortality rate improvement following FCR treatment in FCR responders	0.2000	0.1000	0.3000	Assumed 50% variation
Multiplier for time to new treatment improvement following FCR treatment in FCR responders	0.2000	0.1000	0.3000	Assumed 50% variation
Multiplier for non-response rate improvement following FCR treatment in FCR responders	0.2000	0.1000	0.3000	Assumed 50% variation
Multiplier for mortality rate improvement following BR treatment in FCR responders	0.2000	0.1000	0.3000	Assumed 50% variation
Multiplier for non-response rate improvement following BR treatment in FCR responders	0.2000	0.1000	0.3000	Assumed 50% variation
Multiplier for movement to BSC improvement following BR treatment in FCR responders	0.2000	0.1000	0.3000	Assumed 50% variation

Parameter variation	Value			Justification for low/high values
	Base	Low	High	
Multiplier for P of relapse improvement following BR treatment in FCR responders	0.2000	0.1000	0.3000	Assumed 50% variation
Frequency of complications				
Average number of complications experienced per cycle of first-line FCR treatment	0.349	0.174	0.523	Assumed 50% variation
Average number of complications experienced per cycle of second-line FCR treatment	0.698	0.349	1.047	Assumed 50% variation
Average number of complications experienced per cycle of first-line FCR treatment: FCR responder	0.174	0.087	0.262	Assumed 50% variation
Average number of complications experienced per cycle of first-line BR treatment	0.363	0.181	0.544	Assumed 50% variation
Average number of complications experienced per cycle of second-line BR treatment	0.725	0.363	1.088	Assumed 50% variation
Average number of complications experienced per cycle of second-line BR treatment: FCR responder	0.363	0.181	0.544	Assumed 50% variation
Average number of complications experienced per cycle of first-line Ofa treatment	0.092	0.046	0.138	Assumed 50% variation
Average number of complications experienced per cycle of second- or third-line Ofa treatment	0.184	0.092	0.277	Assumed 50% variation
Average number of complications experienced per cycle of refractory IB treatment	0.058	0.029	0.087	Assumed 50% variation
Average number of complications experienced per cycle of refractory IB treatment: FCR responder	0.029	0.015	0.044	Assumed 50% variation
Average number of complications experienced per cycle of first-line IB treatment: FCR non-responder	0.029	0.015	0.044	Assumed 50% variation
Utilities				
Undergoing first-line treatment	0.803	0.624	0.910	Low value: average of values in Napp <i>et al.</i> [238] and Ferguson <i>et al.</i> [339]. High value: utility assumed to be the same as remission
Undergoing second line treatment	0.710	0.680	0.750	Informed by 95% CI in Beusterien <i>et al.</i> [239]
Undergoing refractory treatment	0.650	0.600	0.690	Informed by 95% CI in Beusterien <i>et al.</i> [239]
Undergoing BMT	0.650	0.585	0.715	Assumed 10% variation
In remission	0.910	0.880	0.930	Informed by 95% CI in Beusterien <i>et al.</i> [239]

Parameter variation	Value			Justification for low/high values
	Base	Low	High	
Average disutility associated with a grade 3/4 adverse event	-0.133	-0.173	-0.100	Informed by 95% CI in Beusterien <i>et al.</i> [239]
Receiving BSC	0.680	0.640	0.720	Informed by 95% CIs in Napp <i>et al.</i> [238]
Other parameters				
Body surface area (m ²)	1.88	1.72	1.93	Informed by estimates used in two NICE submissions [238,315]
Discount rate: costs	0.035	0	0.050	Assumption
Discount rate: outcomes	0.035	0	0.050	Assumption
Number of units of blood received during blood transplant	2.10	1.05	3.15	Assumed 50% variation
Average number of cycles of Ofa treatment	5.76	2.88	8.64	Assumed 50% variation
Average number of cycles of IB treatment	9.35	4.67	14.02	Assumed 50% variation
Average number of cycles spent in remission following FCR treatment	56.31	50.22	62.61	Informed by 95% CI in Hallek <i>et al.</i> [230]
Average number of cycles spent in remission following BMT	51.32	25.66	76.98	Assumed 50% variation
Average number of cycles of FCR treatment (excluding the first cycle)	4.20	2.10	6.30	Assumed 50% variation
Average number of cycles of BR treatment (excluding the first cycle)	3.53	1.76	5.29	Assumed 50% variation
Increase in complications: first-line FCR treatment to second-line FCR treatment (multiplier)	2.00	1.00	3.00	Assumed 50% variation
Decrease in complications: first-line FCR treatment to first-line FCR treatment in FCR responders (multiplier)	0.50	0.25	0.75	Assumed 50% variation
Decrease in complications: second-line BR treatment to first-line BR treatment (multiplier)	0.50	0.25	0.75	Assumed 50% variation
Decrease in complications: second-line BR treatment to second-line BR treatment in FCR responders (multiplier)	0.50	0.25	0.75	Assumed 50% variation
Decrease in complications: second-line Ofa treatment to first-line Ofa treatment (multiplier)	0.50	0.25	0.75	Assumed 50% variation
Decrease in complications: refractory IB treatment to refractory IB treatment in FCR responders (multiplier)	0.50	0.25	0.75	Assumed 50% variation
Decrease in complications: refractory IB treatment to first-line IB treatment in FCR non-responders (multiplier)	0.50	0.25	0.75	Assumed 50% variation

BMT = bone marrow transplant; BR = bendamustine and rituximab; CI = confidence interval; CT = computerised tomography; FBC = full blood count; FCR = fludarabine, cyclophosphamide and rituximab; GCSF = granulocyte colony-stimulating factor; IB = ibrutinib; IgV_H = immunoglobulin variable region; IVIG = IV immunoglobulin; LFT = liver function test; Ofa = ofatumumab; TP = transition probability; U&E = urea and electrolytes.

Appendix 24: Sensitivity analysis results

Parameter variation: The proportion of patients in Comparators A and B who do not have *TP53* mutations – varied between 85% and 95% (base case 93%)

Table 1a: Sensitivity analysis results for this parameter variation – low value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.21	£72,182	DOM	DOM
	Int 2	6.65	£91,790	£81,301	EXT.DOM
	B	7.19	£106,578	£27,035	£45,041
	Int 1	7.45	£119,088	£48,723	£48,723
CUA	C	5.60	£69,704	-	-
	A	5.45	£72,182	DOM	DOM
	Int 2	5.94	£91,790	£66,403	EXT.DOM
	B	6.07	£106,578	£103,233	EXT.DOM
	Int 1	6.68	£119,088	£21,105	£46,213

Table 1b: Sensitivity analysis results for this parameter variation – high value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.70	£71,446	£5,379	£5,379
	Int 2	6.65	£91,790	DOM	DOM
	B	7.73	£107,945	£35,485	£35,485
	Int 1	7.45	£119,088	DOM	DOM
CUA	C	5.60	£69,704	-	-
	A	5.90	£71,446	£5,852	£5,852
	Int 2	5.94	£91,790	£581,258	EXT.DOM
	B	6.52	£107,945	£27,395	£58,425
	Int 1	6.68	£119,088	£76,170	£76,170

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

Cost-effectiveness analysis (CEA) results: When a lower proportion of patients have a *TP53* mutation (as identified by genetic testing), the most cost-effective strategy changes from Comparator A to Comparator C.

Cost-utility analysis (CUA) results: The same effect is observed in the CUA.

Parameter variation: 28-day transition probability of dying from CLL during first-line FCR treatment: no *TP53* mutation – varied between 0.0014 and 0.0041 (base case 0.0027)

Table 2a: Sensitivity analysis results for this parameter variation – low value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	7.07	£74,826	£7,316	£7,316
	Int 2	6.65	£91,790	DOM	DOM
	B	8.16	£113,167	£35,386	£35,386
	Int 1	7.45	£119,088	DOM	DOM
CUA	C	5.60	£69,704	-	-
	A	6.23	£74,826	£8,117	£8,117
	Int 2	5.93	£91,790	DOM	DOM
	B	6.89	£113,167	£57,948	£57,948
	Int 1	6.67	£119,088	DOM	DOM

Table 2b: Sensitivity analysis results for this parameter variation – high value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	A	6.20	£68,693	-	-
	C	6.37	£69,704	£5,855	£5,855
	Int 2	6.65	£91,790	£81,301	EXT.DOM
	B	7.17	£102,824	£21,190	£41,797
	Int 1	7.45	£118,985	£57,462	£57,462
CUA	A	5.45	£68,693	-	-
	C	5.60	£69,704	£6,921	£6,921
	Int 2	5.93	£91,790	£66,403	EXT.DOM
	B	6.04	£102,824	£97,591	EXT.DOM
	Int 1	6.67	£119,088	£26,109	£46,213

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

CEA results: When the probability of dying from chronic lymphocytic leukaemia (CLL) during first-line rituximab, cyclophosphamide and fludarabine (FCR) treatment increases in patients without *TP53* mutations, the most cost-effective strategy changes from Comparator A to Comparator C.

CUA results: The same effect is observed in the CUA.

Parameter variation: 28-day transition probability of moving to second-line treatment when in remission following first-line FCR/bendamustine and rituximab (BR) treatment: no *TP53* mutation – varied between 0.0042 and 0.0126 (base case 0.0084)

Table 3a: Sensitivity analysis results for this parameter variation – low value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	A	7.63	£65,360	-	-
	C	6.37	£69,704	DOM	DOM
	B	8.31	£89,135	£35,108	£35,108
	Int 2	6.65	£91,790	DOM	DOM
	Int 1	7.45	£119,088	DOM	DOM
CUA	A	6.79	£65,360	-	-
	C	5.60	£69,704	DOM	DOM
	B	7.21	£89,135	£56,675	£56,675
	Int 2	5.93	£91,790	DOM	DOM
	Int 1	6.67	£119,088	DOM	DOM

Table 3b: Sensitivity analysis results for this parameter variation – high value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	5.96	£75,004	DOM	DOM
	Int 2	6.65	£91,790	£81,301	EXT.DOM
	B	7.20	£118,899	£49,032	EXT.DOM
	Int 1	7.45	£119,088	£753	£45,920
CUA	C	5.60	£69,704	-	-
	A	5.20	£75,004	DOM	DOM
	Int 2	5.93	£91,790	£66,403	EXT.DOM
	B	5.96	£118,899	£1,021,342	EXT.DOM
	Int 1	6.67	£119,088	£266	£46,213

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

CEA results: When the probability of moving to second-line treatment when in remission following first-line FCR/BR treatment increases in patients without *TP53* mutations, the most cost-effective strategy changes from Comparator A to Comparator C.

CUA results: The same effect is observed in the CUA.

Parameter variation: 28-day transition probability of moving to refractory treatment when in remission following second-line FCR/BR treatment: no *TP53* mutation – varied between 0.0092 and 0.0277 (base case 0.0185)

Table 4a: Sensitivity analysis results for this parameter variation – low value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	A	6.77	£67,806	-	-
	C	6.37	£69,704	DOM	DOM
	Int 2	6.65	£91,790	DOM	DOM
	B	7.47	£92,526	£35,171	£35,171
	Int 1	7.45	£119,088	DOM	DOM
CUA	A	5.98	£67,806	-	-
	C	5.60	£69,704	DOM	DOM
	Int 2	5.93	£91,790	DOM	DOM
	B	6.41	£92,526	£56,801	£56,801
	Int 1	6.67	£119,088	£103,508	£103,508

Table 4b: Sensitivity analysis results for this parameter variation – high value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.51	£73,870	£29,590	£29,590
	Int 2	6.65	£91,790	£136,925	EXT.DOM
	B	7.74	£117,145	£23,269	£35,456
	Int 1	7.45	£119,088	DOM	DOM
CUA	C	5.60	£69,704	-	-
	A	5.72	£73,870	£34,590	£34,590
	Int 2	5.93	£91,790	£84,461	EXT.DOM
	B	6.46	£117,145	£47,689	EXT.DOM
	Int 1	6.67	£119,088	£9,511	£47,690

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

CEA results: Comparator A remains the most cost-effective strategy in both cases.

CUA results: When the probability of moving to refractory treatment when in remission following second-line FCR/BR treatment increases in patients without *TP53* mutations, the most cost-effective strategy changes from Comparator A to Comparator C.

Parameter variation: 28-day transition probability of dying from CLL during first-line FCR treatment: any *TP53* mutation status – varied between 0.0018 and 0.0053 (base case 0.0035)

Table 5a: Sensitivity analysis results for this parameter variation – low value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	A	6.61	£71,576	-	-
	C	6.93	£73,882	£7,245	£7,245
	Int 2	6.65	£91,790	DOM	DOM
	B	7.63	£107,703	£48,114	£48,114
	Int 1	7.45	£119,088	DOM	DOM
CUA	A	5.82	£71,576	-	-
	C	6.09	£73,882	£8,402	£8,402
	Int 2	5.93	£91,790	DOM	DOM
	B	6.44	£107,703	£96,504	EXT.DOM
	Int 1	6.67	£119,088	£50,559	£78,531

Table 5b: Sensitivity analysis results for this parameter variation – high value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	5.90	£66,094	-	-
	A	6.61	£71,576	£7,671	£7,671
	Int 2	6.65	£91,790	£580,390	EXT.DOM
	B	7.63	£107,703	£16,133	£35,376
	Int 1	7.45	£119,088	DOM	DOM
CUA	C	5.17	£66,094	-	-
	A	5.82	£71,576	£8,529	£8,529
	Int 2	5.93	£91,790	£177,198	EXT.DOM
	B	6.44	£107,703	£31,153	EXT.DOM
	Int 1	6.67	£119,088	£50,559	£55,891

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

CEA results: When the probability of dying from CLL during first-line FCR treatment decreases in patients of any *TP53* mutation status, the most cost-effective strategy changes from Comparator A to Comparator C.

CUA results: The same effect is observed in the CUA.

Parameter variation: 28-day transition probability of dying from CLL during second-line BR treatment: any *TP53* mutation status – varied between 0.0093 and 0.0280 (base case 0.0186)

Table 6a: Sensitivity analysis results for this parameter variation – low value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	A	6.61	£71,576	-	-
	C	6.72	£74,265	£25,590	£25,590
	Int 2	6.65	£91,790	DOM	DOM
	B	7.63	£107,703	£36,499	£36,499
	Int 1	7.45	£119,088	DOM	DOM
CUA	A	5.82	£71,576	-	-
	C	5.89	£74,265	£35,518	£35,518
	Int 2	5.93	£91,790	£456,772	EXT.DOM
	B	6.44	£107,703	£31,153	EXT.DOM
	Int 1	6.67	£119,088	£50,559	£57,883

Table 6b: Sensitivity analysis results for this parameter variation – high value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.13	£66,435	-	-
	A	6.61	£71,576	£10,789	£10,789
	Int 2	6.65	£91,790	£580,390	EXT.DOM
	B	7.63	£107,703	£16,133	£35,376
	Int 1	7.45	£119,088	DOM	DOM
CUA	C	5.39	£66,435	-	-
	A	5.82	£71,576	£12,113	£12,113
	Int 2	5.93	£91,790	£177,198	EXT.DOM
	B	6.44	£107,703	£31,153	EXT.DOM
	Int 1	6.67	£119,088	£50,559	£55,891

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

CEA results: When the probability of dying from CLL during second-line BR treatment decreases in patients of any *TP53* mutation status, the most cost-effective strategy changes from Comparator A to Comparator C.

CUA results: Comparator A remains the most cost-effective strategy in both cases.

Parameter variation: 28-day transition probability of moving to second-line treatment when in remission following first-line FCR/BR treatment: any *TP53* mutation status – varied between 0.0046 and 0.0139 (base case 0.0093)

Table 7a: Sensitivity analysis results for this parameter variation – low value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	7.37	£62,791	-	-
	A	6.61	£71,576	DOM	DOM
	Int 2	6.65	£91,790	DOM	DOM
	B	7.63	£107,703	£173,634	£173,634
	Int 1	7.45	£119,088	DOM	DOM
CUA	C	6.55	£62,791	-	-
	A	5.82	£71,576	DOM	DOM
	Int 2	5.93	£91,790	DOM	DOM
	B	6.44	£107,703	DOM	DOM
	Int 1	6.67	£119,088	£483,304	£451,177

Table 7b: Sensitivity analysis results for this parameter variation – high value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	A	6.61	£71,576	-	-
	C	5.75	£73,544	DOM	DOM
	Int 2	6.65	£91,790	£578,264	EXT.DOM
	B	7.63	£107,703	£16,208	£35,376
	Int 1	7.45	£119,088	DOM	DOM
CUA	A	5.82	£71,576	-	-
	C	5.01	£73,544	DOM	DOM
	Int 2	5.93	£91,790	£177,198	EXT.DOM
	B	6.44	£107,703	£31,153	EXT.DOM
	Int 1	6.67	£119,088	£50,559	£55,891

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

CEA results: When the probability of moving to second-line treatment when in remission following first-line FCR/BR treatment decreases in patients of any *TP53* mutation status, the most cost-effective strategy changes from Comparator A to Comparator C.

CUA results: The same effect is observed in the CUA.

Parameter variation: 28-day transition probability of moving to refractory treatment when in remission following second-line FCR/BR treatment: any *TP53* mutation status – varied between 0.0109 and 0.0327 (base case 0.0218)

Table 8a: Sensitivity analysis results for this parameter variation – low value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.53	£65,617	-	-
	A	6.61	£71,576	£77,219	EXT.DOM
	Int 2	6.65	£91,790	£580,390	EXT.DOM
	B	7.63	£107,703	£16,133	£38,316
	Int 1	7.45	£119,088	DOM	DOM
CUA	C	5.76	£65,617	-	-
	A	5.82	£71,576	£110,181	EXT.DOM
	Int 2	5.93	£91,790	£177,198	EXT.DOM
	B	6.44	£107,703	£31,153	EXT.DOM
	Int 1	6.67	£119,088	£50,559	£59,139

Table 8b: Sensitivity analysis results for this parameter variation – high value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	A	6.61	£71,576	-	-
	C	6.28	£72,106	DOM	DOM
	Int 2	6.65	£91,790	£580,390	EXT.DOM
	B	7.63	£107,703	£16,133	£35,376
	Int 1	7.45	£119,088	DOM	DOM
CUA	A	5.82	£71,576	-	-
	C	5.50	£72,106	DOM	DOM
	Int 2	5.93	£91,790	£177,198	EXT.DOM
	B	6.44	£107,703	£31,153	EXT.DOM
	Int 1	6.67	£119,088	£50,559	£55,891

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

CEA results: When the probability of moving to refractory treatment when in remission following second-line FCR/BR treatment decreases in patients of any *TP53* mutation status, the most cost-effective strategy changes from Comparator A to Comparator C.

CUA results: The same effect is observed in the CUA.

Parameter variation: 28-day transition probability of moving to second-line treatment when in remission following first-line FCR treatment: FCR responder, correctly identified – varied between 0.0035 and 0.0105 (base case 0.0070)

Table 9a: Sensitivity analysis results for this parameter variation – low value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.61	£71,576	£7,903	£7,903
	Int 2	7.12	£82,328	£21,053	£21,053
	B	7.63	£107,703	£49,705	EXT.DOM
	Int 1	7.93	£109,627	£6,559	£33,963
CUA	C	5.60	£69,704	-	-
	A	5.82	£71,576	£8,565	£8,565
	Int 2	6.38	£82,328	£19,248	£19,248
	B	6.44	£107,703	£382,788	EXT.DOM
	Int 1	7.11	£109,627	£2,872	£37,090

Table 9b: Sensitivity analysis results for this parameter variation – high value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.61	£71,576	£7,903	£7,903
	Int 2	6.33	£97,755	DOM	DOM
	B	7.63	£107,703	£35,376	£35,376
	Int 1	7.13	£125,054	DOM	DOM
CUA	C	5.60	£69,704	-	-
	A	5.82	£71,576	£8,565	£8,565
	Int 2	5.63	£97,755	DOM	DOM
	B	6.44	£107,703	£57,813	£57,813
	Int 1	6.37	£125,054	DOM	DOM

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

CEA results: When the probability of moving to second-line treatment when in remission following first-line FCR treatment decreases in correctly identified FCR responders, the most cost-effective strategy changes from Comparator A to Intervention Two.

CUA results: The same effect is observed in the CUA.

Parameter variation: Cost discount rate – varied between 0% and 5% (base case 3.5%)

Table 10a: Sensitivity analysis results for this parameter variation – low value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£79,076	-	-
	A	6.61	£81,373	£9,698	£9,698
	Int 2	6.65	£114,074	£938,954	EXT.DOM
	B	7.63	£137,738	£23,990	£55,194
	Int 1	7.45	£150,823	DOM	DOM
CUA	C	5.60	£79,076	-	-
	A	5.82	£81,373	£10,511	£10,511
	Int 2	5.93	£114,074	£286,670	EXT.DOM
	B	6.44	£137,738	£46,324	EXT.DOM
	Int 1	6.67	£150,823	£58,108	£81,698

Table 10b: Sensitivity analysis results for this parameter variation – high value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£66,567	-	-
	A	6.61	£68,341	£7,493	£7,493
	Int 2	6.65	£85,034	£479,305	EXT.DOM
	B	7.63	£98,637	£13,790	£29,666
	Int 1	7.45	£109,366	DOM	DOM
CUA	C	5.60	£66,567	-	-
	A	5.82	£68,341	£8,120	£8,120
	Int 2	5.93	£85,034	£146,336	EXT.DOM
	B	6.44	£98,637	£26,628	EXT.DOM
	Int 1	6.67	£109,366	£47,647	£48,260

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

CEA results: When the rate used to discount costs in the model increases to 5%, the most cost-effective strategy changes from Comparator A to Comparator B.

CUA results: Comparator A remains the most cost-effective strategy in both cases.

Parameter variation: Outcome discount rate – varied between 0% and 5% (base case 3.5%)

Table 11a: Sensitivity analysis results for this parameter variation – low value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	7.51	£69,704	-	-
	A	7.89	£71,576	£4,944	£4,944
	Int 2	8.21	£91,790	£62,725	EXT.DOM
	B	9.46	£107,703	£12,724	£22,968
	Int 1	9.28	£119,088	DOM	DOM
CUA	C	6.60	£69,704	-	-
	A	6.95	£71,576	£5,389	£5,389
	Int 2	7.34	£91,790	£51,579	EXT.DOM
	B	7.91	£107,703	£27,794	EXT.DOM
	Int 1	8.32	£119,088	£28,388	£34,795

Table 11b: Sensitivity analysis results for this parameter variation – high value

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	5.99	£69,704	-	-
	A	6.19	£71,576	£9,580	£9,580
	Int 2	6.15	£91,790	DOM	DOM
	B	7.05	£107,703	£41,901	£41,901
	Int 1	6.87	£119,088	DOM	DOM
CUA	C	5.26	£69,704	-	-
	A	5.44	£71,576	£10,355	£10,355
	Int 2	5.48	£91,790	£482,101	EXT.DOM
	B	5.97	£107,703	£32,781	EXT.DOM
	Int 1	6.15	£119,088	£64,566	£67,517

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

CEA results: When the rate used to discount outcomes in the model decreases to 0%, the most cost-effective strategy changes from Comparator A to Comparator B.

CUA results: Comparator A remains the most cost-effective strategy in both cases.

Appendix 25: Scenario analysis results

Scenario One: Changes in the cost of Ibrutinib treatment

Table 1a: 50% reduction

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.61	£71,576	£7,903	£7,903
	Int 2	6.65	£74,460	£82,818	EXT.DOM
	Int 1	7.45	£86,992	£15,592	EXT.DOM
	B	7.63	£87,027	£188	£15,130
CUA	C	5.60	£69,704	-	-
	A	5.82	£71,576	£8,565	£8,565
	Int 2	5.93	£74,460	£25,285	EXT.DOM
	Int 1	6.67	£86,992	£17,027	£18,135
	B	6.44	£87,027	DOM	DOM

Table 1b: 67% reduction

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	Int 2	6.65	£68,684	-	-
	C	6.37	£69,704	DOM	DOM
	A	6.61	£71,576	DOM	DOM
	Int 1	7.45	£76,294	£9,468	£9,468
	B	7.63	£80,135	£21,032	£21,032
CUA	Int 2	5.93	£68,684	-	-
	C	5.60	£69,704	DOM	DOM
	A	5.82	£71,576	DOM	DOM
	Int 1	6.67	£76,294	£10,340	£10,340
	B	6.44	£80,135	DOM	DOM

Table 1c: 75% reduction

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	Int 2	6.65	£65,796	-	-
	C	6.37	£69,704	DOM	DOM
	Int 1	7.45	£70,945	£6,406	£6,406
	A	6.61	£71,576	DOM	DOM
	B	7.63	£76,689	£31,453	£31,453
CUA	Int 2	5.93	£65,796	-	-
	C	5.60	£69,704	DOM	DOM
	Int 1	6.67	£70,945	£6,996	£6,996
	A	5.82	£71,576	DOM	DOM
	B	6.44	£76,689	DOM	DOM

Table 1d: increase to £63,011 per year (to match the cost in the USA)

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.61	£71,576	£7,903	£7,903
	Int 2	6.65	£129,927	£1,675,410	EXT.DOM
	B	7.63	£153,207	£23,601	£79,934
	Int 1	7.45	£189,722	DOM	DOM
CUA	C	5.60	£69,704	-	-
	A	5.82	£71,576	£8,565	£8,565
	Int 2	5.93	£129,927	£511,515	EXT.DOM
	B	6.44	£153,207	£45,573	£130,629
	Int 1	6.67	£189,722	£162,160	£162,160

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

Comments: Comparator A is no longer the most cost-effective strategy in any scenario in which the cost of ibrutinib treatment is reduced. Instead, one of the strategies which uses ibrutinib treatment is always the most cost-effective option. In the cost-utility analysis (CUA), the most cost-effective strategy is always Intervention One. In the cost-effectiveness analysis (CEA), the most cost-effective strategy is Comparator B for 50% and 67% reductions in the cost of ibrutinib. For a 75% reduction in the cost, the most cost-effective strategy is Intervention One. An increase in the cost of ibrutinib to match the cost of treatment in the USA does not change the results, with Comparator A remaining the most cost-effective strategy in both the CEA and the CUA.

Scenario Two: No genetic or genomic testing before subsequent lines of treatment

Table 2: No genetic or genomic testing before subsequent lines of treatment

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.61	£71,358	£6,984	£6,984
	Int 2	6.65	£91,226	£570,451	EXT.DOM
	B	7.63	£107,021	£16,013	£34,921
	Int 1	7.45	£118,524	DOM	DOM
CUA	C	5.60	£69,704	-	-
	A	5.82	£71,358	£7,568	£7,568
	Int 2	5.93	£91,226	£174,163	EXT.DOM
	B	6.44	£107,021	£30,921	EXT.DOM
	Int 1	6.67	£118,524	£51,085	£55,484

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

Comments: Comparator A remains the most cost-effective strategy.

Scenario Three: Variations in the improved treatment effect applied following genomic testing

Table 3a: Assumed improved treatment effect is 40% rather than 20%

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.61	£71,576	£7,903	£7,903
	Int 2	7.26	£88,693	£26,264	£26,264
	B	7.63	£107,703	£51,449	EXT.DOM
	Int 1	8.07	£115,992	£19,085	£33,963
CUA	C	5.60	£69,704	-	-
	A	5.82	£71,576	£8,565	£8,565
	Int 2	6.50	£88,693	£25,245	£25,245
	B	6.44	£107,703	DOM	DOM
	Int 1	7.23	£115,992	£37,090	£37,090

Table 3b: Assumed improved treatment effect is 60% rather than 20%

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.61	£71,576	£7,903	£7,903
	Int 2	8.02	£84,814	£9,372	£9,372
	B	7.63	£107,703	DOM	DOM
	Int 1	8.83	£112,113	£33,963	£33,963
CUA	C	5.60	£69,704	-	-
	A	5.82	£71,576	£8,565	£8,565
	Int 2	7.19	£84,814	£9,640	£9,640
	B	6.44	£107,703	DOM	DOM
	Int 1	7.93	£112,113	£37,090	£37,090

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

Comments: In both cases, the most cost-effective strategy changes from Comparator A to Intervention Two.

Scenario Four: Variations in patient starting ages

Table 4a: Starting age is 25

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	7.00	£74,899	-	-
	A	7.35	£77,313	£6,829	£6,829
	Int 2	7.73	£109,895	£85,369	EXT.DOM
	B	8.88	£131,996	£19,269	£35,772
	Int 1	8.73	£143,896	DOM	DOM
CUA	C	6.15	£74,899	-	-
	A	6.47	£77,313	£7,458	£7,458
	Int 2	6.91	£109,895	£74,521	EXT.DOM
	B	7.42	£131,996	£43,872	EXT.DOM
	Int 1	7.82	£143,896	£29,570	£49,563

Table 4b: Starting age is 85

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	3.90	£49,424	-	-
	A	3.91	£50,954	£82,931	EXT.DOM
	Int 2	3.65	£53,477	DOM	DOM
	B	4.11	£56,874	£29,947	£34,472
	Int 1	3.97	£63,354	DOM	DOM
CUA	C	3.41	£49,424	-	-
	A	3.43	£50,954	£82,011	EXT.DOM
	Int 2	3.23	£53,477	DOM	DOM
	B	3.55	£56,874	£48,416	£48,416
	Int 1	3.54	£63,354	DOM	DOM

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

Comments: Comparator A remains the most cost-effective strategy for patients aged 25. However, for patients aged 85, the most cost-effective strategy changes to Comparator C.

Scenario Five: Information on the presence of *SAMHD1* mutations excluded from the analysis

Table 5: *SAMHD1* data excluded

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.61	£71,576	£7,903	£7,903
	Int 2	6.79	£91,419	£109,937	EXT.DOM
	B	7.63	£107,703	£19,370	£35,376
	Int 1	7.34	£110,031	DOM	DOM
CUA	C	5.60	£69,704	-	-
	A	5.82	£71,576	£8,565	£8,565
	Int 2	6.07	£91,419	£79,993	EXT.DOM
	B	6.44	£107,703	£43,213	EXT.DOM
	Int 1	6.57	£110,031	£18,627	£51,282

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

Comments: Comparator A remains the most cost-effective strategy.

Scenario Six: Ibrutinib treatment in widespread use

Table 6: Ibrutinib treatment in widespread use

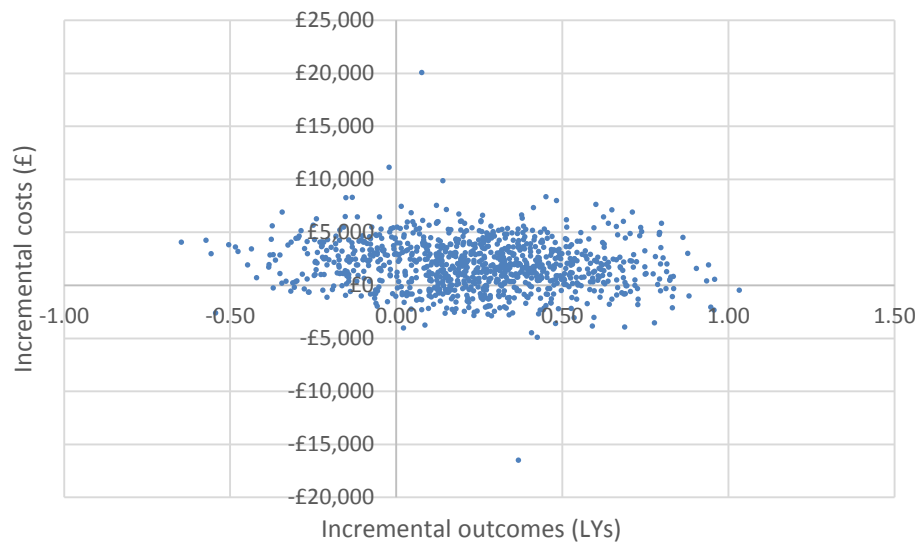
Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	Int 2	6.65	£91,790	-	-
	B	7.63	£107,703	£16,133	£16,133
	Int 1	7.45	£119,088	DOM	DOM
CUA	Int 2	5.93	£91,790	-	-
	B	6.44	£107,703	£31,153	£31,153
	Int 1	6.67	£119,088	£50,559	£50,559

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life years; QALYs = quality-adjusted life years.

Comments: The most cost-effective strategy is either Comparator B (CEA) or Intervention Two (CUA).

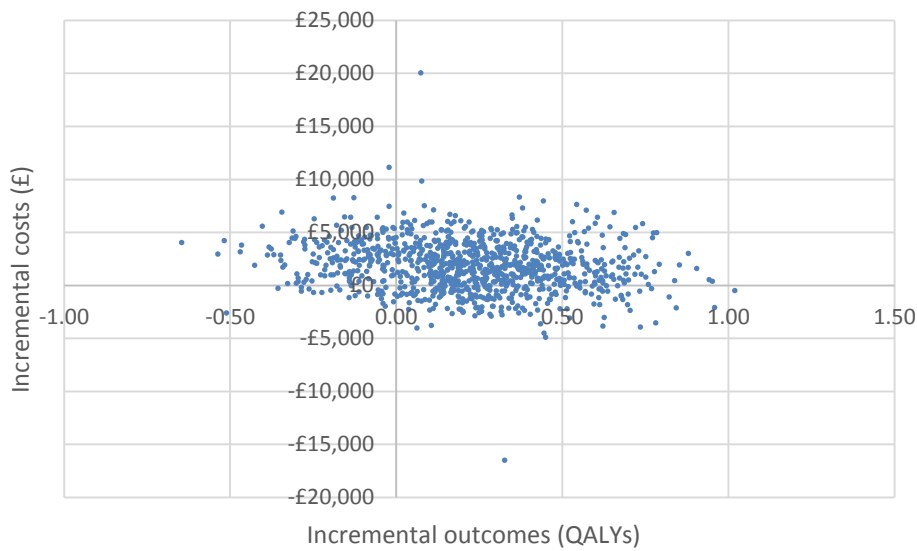
Appendix 26: Cost-effectiveness planes

Figure 1a: Cost and effect pairs generated by the probabilistic sensitivity analysis for the comparison between Comparator A and Comparator C in the cost-effectiveness analysis



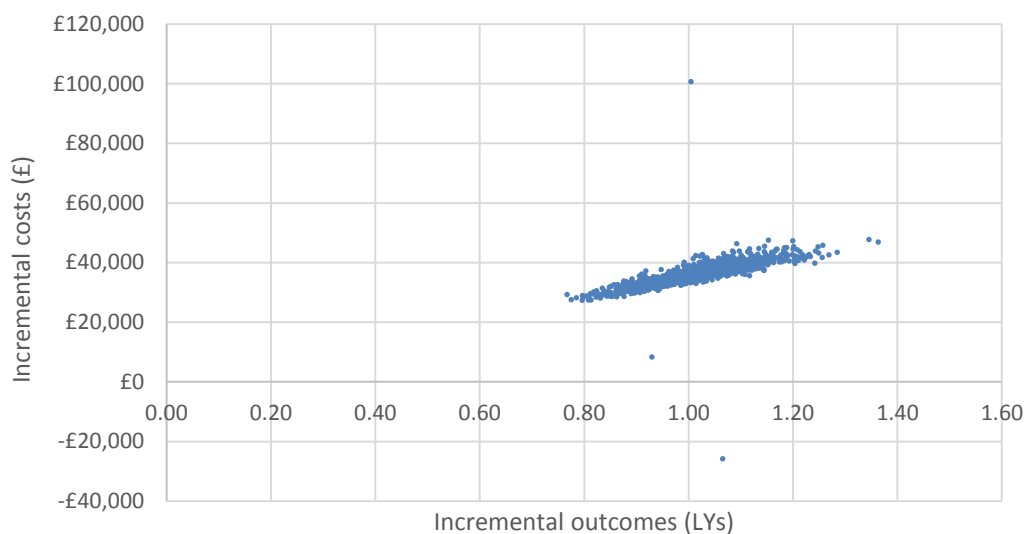
Note: these are the incremental costs and outcomes for Comparator A versus Comparator C.

Figure 1b: Cost and effect pairs generated by the probabilistic sensitivity analysis for the comparison between Comparator A and Comparator C in the cost-utility analysis



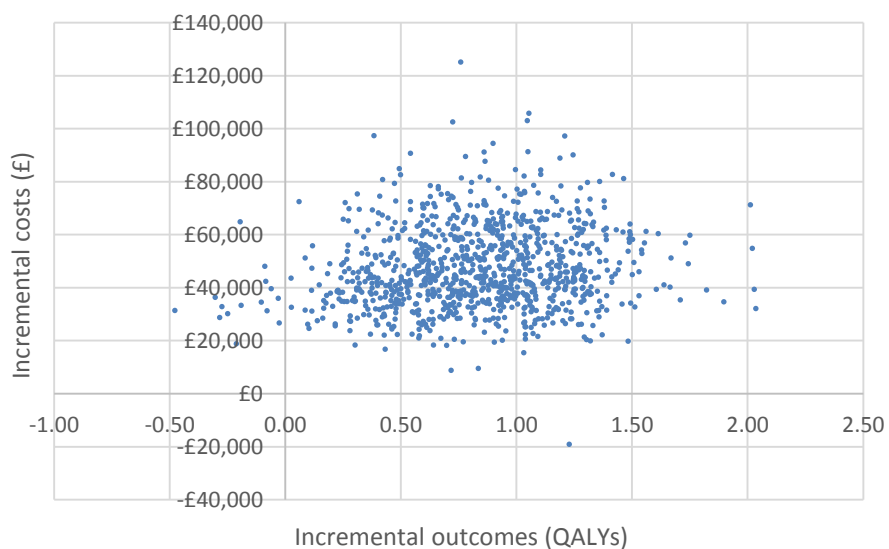
Note: these are the incremental costs and outcomes for Comparator A versus Comparator C.

Figure 2: Cost and effect pairs generated by the probabilistic sensitivity analysis for the comparison between Comparator B and Comparator A in the cost-effectiveness analysis



Note: these are the incremental costs and outcomes for Comparator B versus Comparator A.

Figure 3: Cost and effect pairs generated by the probabilistic sensitivity analysis for the comparison between Intervention One and Comparator A in the cost-utility analysis



Note: these are the incremental costs and outcomes for Intervention One versus Comparator A.

Appendix 27: Literature search for discrete choice experiments conducted in genetics or leukaemia

Background

The aim of this exploratory literature search was to identify discrete choice experiments (DCEs) that had been conducted in either a genetic context (including genomic testing) or in any form of leukaemia, as well as articles discussing the use of DCEs in either of these two contexts from a methodological perspective. References identified in this review would then be used to inform the design of the DCE presented in this chapter. For this reason, broad searches were undertaken. These searches were not limited to full length references as it was acknowledged earlier in this chapter that it is not always straightforward to identify DCEs in the published literature, so references such as conference abstracts and posters were also sought. Literature searches were undertaken in Embase, Medline, Econlit, the Allied and Complementary Medicine database, CAB abstracts, the Global Health database, the UK National Health Service economic evaluation database (NHS EED) and UK NHS Health Technology Assessment programme (both via the Centre for Reviews and Dissemination (CRD) database), and The Cochrane Library. The initial search was undertaken in November 2011 and updated in September 2012.

Search strategies

Ovid

The Ovid interface was used to search the following databases: Embase, Medline, Econlit, the Allied and Complementary Medicine database, CAB abstracts and the Global Health database. The following syntax was used across all databases:

```
genetic* or mutation* or deletion* or duplication* or chromosomal imbalance or genomic*  
leukaemia or leukemia or haematological or B-CLL or chronic lymphocytic  
discrete choice or DCE or conjoint or stated preference  
1 and 3  
2 and 3  
4 or 5  
limit 6 to english language  
limit 7 to yr="1990-current"  
remove duplicates from 8
```

CRD

discrete choice or DCE or conjoint or stated preference

The Cochrane Library

discrete choice or DCE or conjoint or stated preference

Results

After reviewing titles and abstracts and removing duplicates, 29 references were identified which were potentially informative. As the literature searches were broad, a number of these references

were not full length articles and were therefore not publicly available. The authors of these articles were contacted and these materials were requested. As some authors did not respond, the final number of references reviewed was 25 [15,147,256-258,271,286,287,341-357]. In some cases, the survey instruments used in these studies were provided as supplementary materials. Where this was not the case, these survey instruments were requested from the study authors. Some studies were described as DCEs but on further investigation were misclassified. In some cases these studies were retained if they contained material that could usefully inform the development of this DCE.

Most of the references (23/25) were published after 2004. The disparate nature of the different studies meant that it was not possible to draw out common findings. However, a variety of useful information was provided that was used to inform the development of the DCE presented in Chapter Five, including:

- Suggestions of background questions that could place responses to the DCE presented in this chapter into context;
- Examples of good practice in terms of phrasing DCE questions and providing appropriate background materials;
- Examples of effective approaches to presenting risk information;
- Suggestions of attributes (and levels for these attributes) that might be important in the context of genomic testing in chronic lymphocytic leukaemia;
- Examples of different payment vehicles that were acceptable to patients in this context;
- Information on potential analytical approaches that are likely to be appropriate for DCEs in genomics;
- Estimates of willingness-to-pay for genetic or genomic testing, or for leukaemia interventions, that could be used as comparators for the results of this DCE.

Appendix 28: Materials used in patient interviews

Answer sheet and script

INTERVIEW NUMBER: _____ **DATE:** _____ **TIME:** _____

PARTICIPATION REQUEST

Hi [PATIENT NAME], my name is James Buchanan. [CLINICIAN NAME] is not ready to see you yet, however I was wondering if you would be willing to help with a project on CLL and genetics that Dr Anna Schuh is running right now, while you wait for your appointment? I would need to ask you a few simple questions about the current understanding of genetic testing within the CLL patient community. We would do this in [INSERT LOCATION HERE] - it would take around 20 minutes of your time and you will not be late for your appointment.

>>> *Would you be willing to help me with this?*

INTRODUCTION

Firstly, thank you for agreeing to participate. I am an Oxford University PhD student working with Dr Anna Schuh on a project which is investigating whether patients with CLL value the opportunity to undergo genetic testing. Several different genetic tests exist for CLL: some tests are more reliable than others, and each provides different information. Sometimes this information can be very useful to clinicians, for example information on whether or not a patient is likely to respond to treatment.

These tests are not yet in widespread use. One of the reasons for this is that we don't know whether CLL patients themselves would value the opportunity to undergo genetic testing. In the next few months I will be conducting a survey amongst CLL patients in Oxford and the rest of the UK which will ask whether they would value having access to genetic testing, and which characteristics (or features) of testing are most important to them, for example, how quickly you get your results back, or how accurate these results are.

Before I conduct this survey, I need to make sure that I'm asking the right questions. In particular I want to make sure that I am asking about the attributes of testing that are the most important to patients. I also want to make sure that patients will understand the questions that I ask in the survey. So I'm conducting several informal interviews with CLL patients here in Oxford to try to answer these initial questions.

>>> *Are you still happy to proceed?*

QUESTIONS

Before I begin, I just want to mention that there are no correct answers to any of these questions. We are only interested in learning more about the current understanding of genetic testing within the CLL patient community, and finding out what characteristics of genetic testing you think are important. All of your answers are completely anonymous and will not affect your current treatment.

[IF PATIENT OBSERVER IS PRESENT] I am happy for your observer to be present, but we are only interested in the thoughts and opinions of current CLL patients for the purposes of this survey, so all of your answers must be your own.

Section A

There are four sections in this interview. In this first section, I would like to ask some questions about your previous experience of genetic testing.

1) You have probably had several different tests since being diagnosed with CLL. Do you know if any of these tests were genetic tests?

☐ Yes ☐ No ☐ Not sure

2) If yes, do you know what sort of genetic tests these were?

3) Did the clinician who organised the tests explain to you why the tests might be useful?

☐ Yes ☐ No ☐ Not sure

4) Did the clinician explain the test results to you when you received them?

☐ Yes ☐ No ☐ Not sure

5) If you were offered a genetic test but decided not to have it, could I ask you why?

Section B

1) In this second section I would like to ask you some questions about the characteristics of different genetic tests. Imagine that you have been offered a genetic test that can provide your clinician with some information that they can use to manage your illness. Specifically, the test will tell you and your clinician whether or not you will respond to the usual chemotherapy treatment that is offered to CLL patients when they are first diagnosed.

You have to decide whether or not you want your clinician to carry out this test. One of the ways in which you might make this decision is by considering the different characteristics of the test and deciding which ones are important to you. I'm going to show you a list of the different characteristics that a genetic test might have, and I will explain each characteristic to you in turn. These characteristics are presented in the order in which you might think about them if you were deciding whether to undergo testing or not. For each characteristic in this list I want you to indicate how important this characteristic would be if you were making a decision to undergo genetic testing. **Normally patients report that some characteristics are much more important to them than others.**

The ranking scale underneath each characteristic runs from 0 to 10. 0 means not at all important and 10 means extremely important. For each attribute, please circle how important it is.

[GIVE LIST TO PATIENTS. GO THROUGH EACH CHARACTERISTIC ONE BY ONE DESCRIBING WHAT THE CHARACTERISTIC IS, THEN ASK THEM TO RATE EACH ONE]

DESCRIPTION OF ATTRIBUTES FOR PATIENTS

1. Level of information given to patients about the test

This is the amount of information that you are given before you have the test. This can include speaking to the person who recommends the test, opportunities to ask questions, and written material you may be given to take home with you. Some people like to have as much information as possible before deciding whether to undergo testing; others prefer to have less.

2. Number of times that a blood sample is required

Each time that you have a genetic test you need to supply a blood sample. Some genetic tests only need to be given once, before the start of treatment, so only one blood sample is needed. Other tests can be given more often, for example before the start of treatment, during treatment, and once your course of treatment has ended. These genetic tests therefore require more blood samples.

3. Location of test

This describes the location where you are given your test results. It could be your local GP surgery, a specialist CLL clinic, or within a general hospital setting e.g. in an outpatient clinic for all patients.

4. Speed of test result

This describes how many days are between giving a blood sample and receiving the test result. You will only start the chemotherapy treatment that you are being tested for once you receive the result, and only if that result indicates that you will respond to treatment.

5. Test effectiveness

A certain percentage of CLL patients will not respond to the usual chemotherapy treatment. The most effective genetic tests can detect a lot of these patients, while less effective tests can only detect a few of these patients.

6. Test reliability

Occasionally a genetic test will not be correct, and will indicate that you will respond to the usual chemotherapy treatment when in fact you won't. The most reliable genetic tests rarely return an incorrect result, while less reliable tests return incorrect results more often.

7. Provision of exact diagnosis

Most patients with CLL receive a correct diagnosis as a result of the different genetic tests that they undergo. However a few patients receive an incorrect diagnosis, for example they are told that they have a pre-leukaemic condition instead of CLL. Some genetic tests are available which increase the number of patients who receive a correct diagnosis.

8. Who explains the test result to you

This describes which person, working for the NHS, explains the test result to you. It could be your GP, or a hospital doctor, or a specialist nurse, for example.

9. Cost of test

Obviously in the NHS you will not be asked to pay any money for genetic testing. However, imagine that you lived in a country where you did have to pay for a genetic test, such as the USA. In that situation, how important would the cost of the test be when you make your decision?

2) This list may not include all of the test characteristics that could influence your decision to undergo genetic testing. Are there any other characteristics that would be important in influencing your decision?

3) Which of these test characteristics would be the most important characteristic for you?

Section C

In this third section I would like to ask you some questions about interpreting risk information. We know that some patients find it difficult to understand information about risk when this is presented to them by a clinician. This information could be described to you verbally, or given as a probability (e.g. 20% risk of an event), or as a frequency (e.g. this event will happen to 20 out of every 100 patients) or described using a diagram (e.g. a picture of 100 figures with 20 shaded in to indicate those will undergo the event).

1) I'm now going to show you 10 examples of different ways that risk information could be presented. In each example you will see a graphic which describes what proportion of CLL patients will respond to a hypothetical chemotherapy treatment. I would like you to tell me what the proportion of patients is in each case. You will have four options to choose from. Take as much time as you need for each question.

[SHOW PATIENTS THE TEN POWERPOINT SLIDES IN SEQUENCE AND ASK THEM TO ANSWER THESE QUESTIONS]

2) I just presented risk information to you using 5 different types of graphic: complex stick figures, a grid of squares, simple stick figures, rectangles, and faces.

(a) Which of these graphics made it easiest for you to understand the risk information?

- ☐ Complex stick figures
- ☐ Grid of squares
- ☐ Simple stick figures
- ☐ Rectangles
- ☐ Faces
- ☐ None of these

(b) Which of these graphics made it hardest for you to understand the risk information?

- ☐ Complex stick figures
- ☐ Grid of squares
- ☐ Simple stick figures
- ☐ Rectangles
- ☐ Faces
- ☐ None of these

Section D

Finally I would like to ask a few questions about your personal details. You do not have to answer these questions but doing so will allow me to see how representative the responses I receive in these interviews are. It is also useful for me to be able to see whether different types of people give different responses. All the details that you provide will be kept confidential.

1) How long ago were you diagnosed with CLL, to the nearest year?

2) Are you currently undergoing any sort of treatment for your CLL?

- ☐ Yes ☐ No

3) If yes, what type of treatment are you currently receiving?

4) If this is chemotherapy, what type of chemotherapy is this?

5) How many rounds of chemotherapy treatment have you had?

6) [Note whether they are male or female]

☐ Male ☐ Female

7) Which of these age bands does your age fall into?

☐ 20-29 years	☐ 60-69 years
☐ 30-39 years	☐ 70-79 years
☐ 40-49 years	☐ 80-89 years
☐ 50-59 years	☐ >90 years

8) How old were you when you left full time education?

9) Are you currently employed?

☐ Yes ☐ No

10) How many children do you have?

11) How long did it take you to get to hospital today?

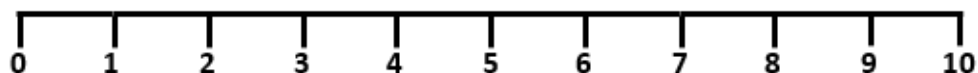
☐ 0-15 minutes	☐ 61-75 minutes
☐ 16-30 minutes	☐ 76-90 minutes
☐ 31-45 minutes	☐ >90 minutes
☐ 46-60 minutes	

Section B for patients

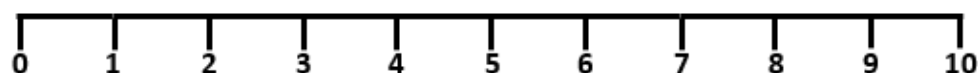
INTERVIEW NUMBER: _____ DATE: _____ TIME: _____

SECTION B

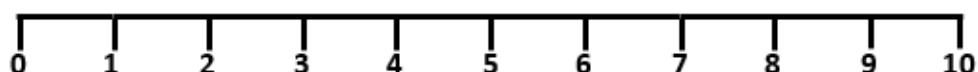
1. LEVEL OF INFORMATION GIVEN TO PATIENTS ABOUT THE TEST



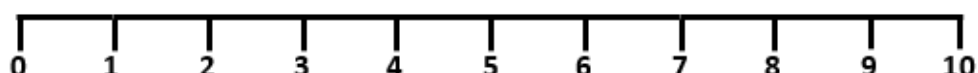
2. NUMBER OF TIMES THAT A BLOOD SAMPLE IS REQUIRED



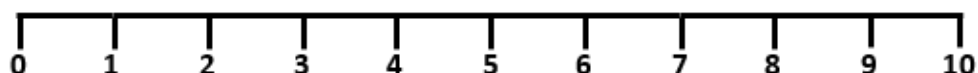
3. LOCATION OF TEST



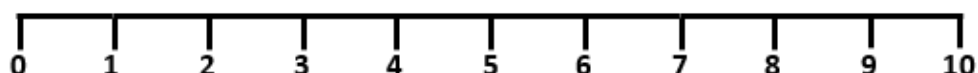
4. SPEED OF TEST RESULT



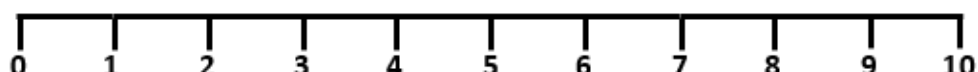
5. TEST EFFECTIVENESS



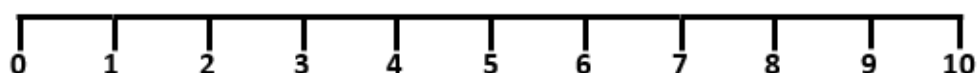
6. TEST RELIABILITY



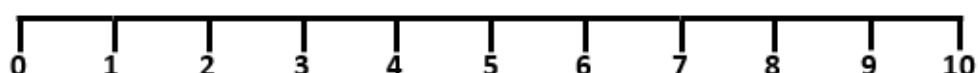
7. PROVISION OF EXACT DIAGNOSIS



8. WHO EXPLAINS THE TEST RESULT TO YOU



9. COST OF TEST



PowerPoint slides shown to patients

1.



Each figure represents one CLL patient.

How many CLL patients will respond to the hypothetical chemotherapy treatment?

Please select 1 of the following 4 options:

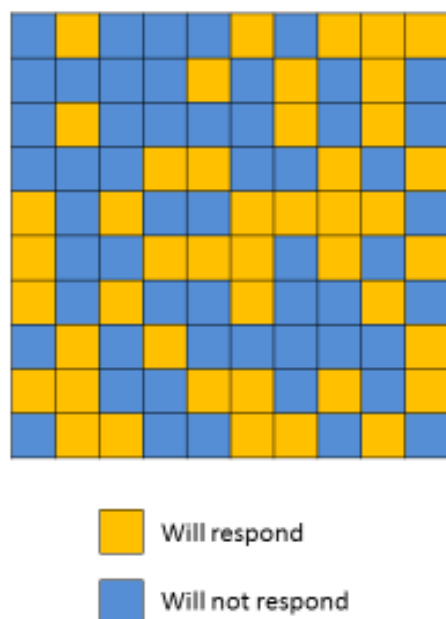
A. 55 out of every 100 patients will respond to chemotherapy ☐

B. 60 out of every 100 patients will respond to chemotherapy ☐

C. 65 out of every 100 patients will respond to chemotherapy ☐

D. 70 out of every 100 patients will respond to chemotherapy ☐

2.



Each square represents one CLL patient.

How many CLL patients will respond to the hypothetical chemotherapy treatment?

Please select 1 of the following 4 options:

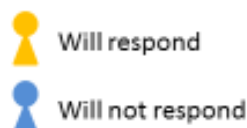
A. 35 out of every 100 patients will respond to chemotherapy ☐

B. 40 out of every 100 patients will respond to chemotherapy ☐

C. 45 out of every 100 patients will respond to chemotherapy ☐

D. 50 out of every 100 patients will respond to chemotherapy ☐

3.



Each figure represents one CLL patient.

How many CLL patients will respond to the hypothetical chemotherapy treatment?

Please select 1 of the following 4 options:

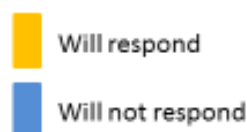
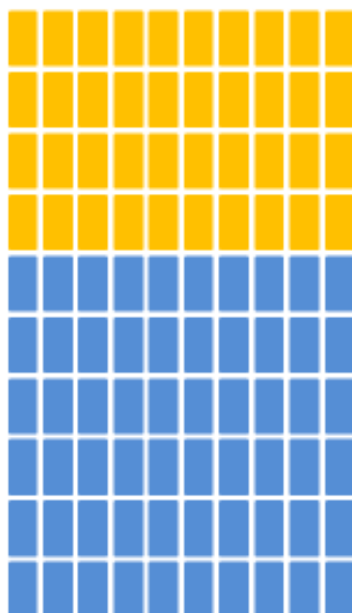
A. 50 out of every 100 patients will respond to chemotherapy ☐

B. 55 out of every 100 patients will respond to chemotherapy ☐

C. 60 out of every 100 patients will respond to chemotherapy ☐

D. 65 out of every 100 patients will respond to chemotherapy ☐

4.



Each block represents one CLL patient.

How many CLL patients will respond to the hypothetical chemotherapy treatment?

Please select 1 of the following 4 options:

A. 35 out of every 100 patients will respond to chemotherapy ☐

B. 40 out of every 100 patients will respond to chemotherapy ☐

C. 45 out of every 100 patients will respond to chemotherapy ☐

D. 50 out of every 100 patients will respond to chemotherapy ☐

5.



 Will respond
 Will not respond

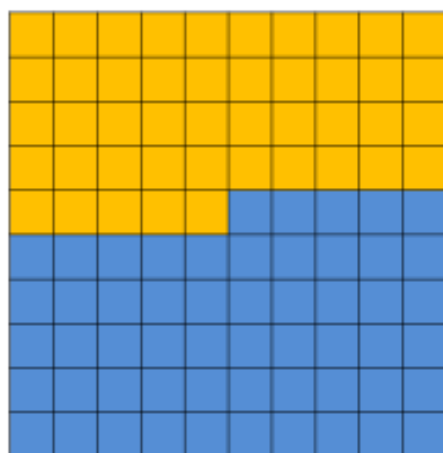
Each face represents one CLL patient.

How many CLL patients will respond to the hypothetical chemotherapy treatment?

Please select 1 of the following 4 options:

- A. 45 out of every 100 patients will respond to chemotherapy ☐
- B. 50 out of every 100 patients will respond to chemotherapy ☐
- C. 55 out of every 100 patients will respond to chemotherapy ☐
- D. 60 out of every 100 patients will respond to chemotherapy ☐

6.



 Will respond
 Will not respond

Each square represents one CLL patient.

How many CLL patients will respond to the hypothetical chemotherapy treatment?

Please select 1 of the following 4 options:

- A. 30 out of every 100 patients will respond to chemotherapy ☐
- B. 35 out of every 100 patients will respond to chemotherapy ☐
- C. 40 out of every 100 patients will respond to chemotherapy ☐
- D. 45 out of every 100 patients will respond to chemotherapy ☐

7.



Will respond



Will not respond

Each figure represents one CLL patient.

How many CLL patients will respond to the hypothetical chemotherapy treatment?

Please select 1 of the following 4 options:

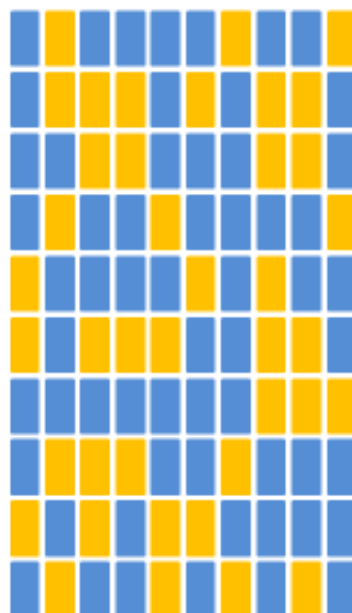
A. 50 out of every 100 patients will respond to chemotherapy ☐

B. 55 out of every 100 patients will respond to chemotherapy ☐

C. 60 out of every 100 patients will respond to chemotherapy ☐

D. 65 out of every 100 patients will respond to chemotherapy ☐

8.



Will respond



Will not respond

Each block represents one CLL patient.

How many CLL patients will respond to the hypothetical chemotherapy treatment?

Please select 1 of the following 4 options:

A. 40 out of every 100 patients will respond to chemotherapy ☐

B. 45 out of every 100 patients will respond to chemotherapy ☐

C. 50 out of every 100 patients will respond to chemotherapy ☐

D. 55 out of every 100 patients will respond to chemotherapy ☐

9.



 Will respond
 Will not respond

Each figure represents one CLL patient.

How many CLL patients will respond to the hypothetical chemotherapy treatment?

Please select 1 of the following 4 options:

A. 45 out of every 100 patients will respond to chemotherapy ☐

B. 50 out of every 100 patients will respond to chemotherapy ☐

C. 55 out of every 100 patients will respond to chemotherapy ☐

D. 60 out of every 100 patients will respond to chemotherapy ☐

10.



 Will respond
 Will not respond

Each face represents one CLL patient.

How many CLL patients will respond to the hypothetical chemotherapy treatment?

Please select 1 of the following 4 options:

A. 40 out of every 100 patients will respond to chemotherapy ☐

B. 45 out of every 100 patients will respond to chemotherapy ☐

C. 50 out of every 100 patients will respond to chemotherapy ☐

D. 55 out of every 100 patients will respond to chemotherapy ☐

Appendix 29: Patient interview methods and results

Objectives

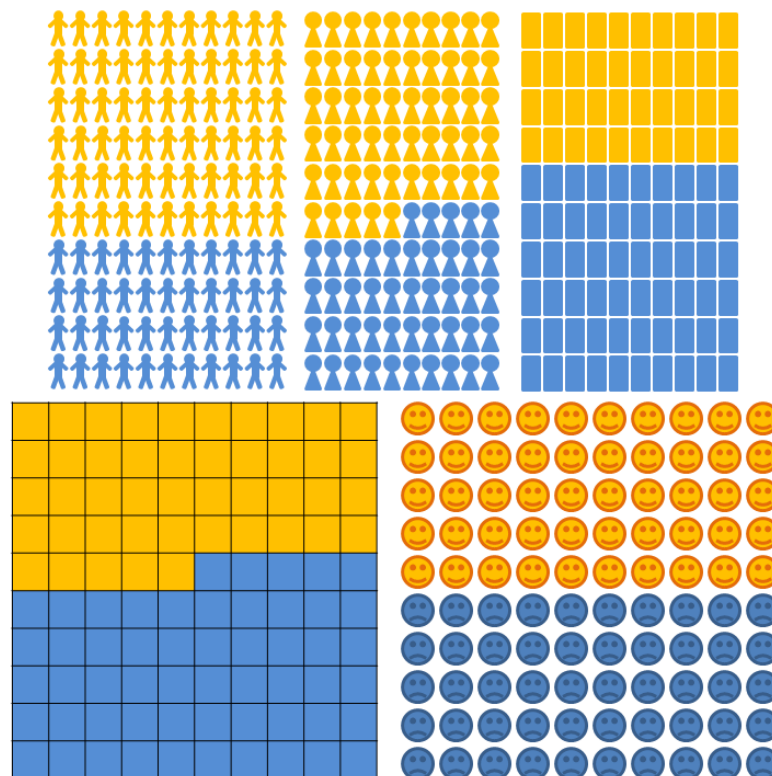
The patient interviews had two objectives:

- To identify which attributes of genetic and genomic tests were most important to chronic lymphocytic leukaemia (CLL) patients;
- To identify the most appropriate graphical approach to use when presenting risk information to patients in the final discrete choice experiment (DCE).

Methods

Interviews were conducted with randomly selected patients attending a CLL outpatient clinic at the Oxford University Hospitals (OUH) NHS Trust. Interviews were held in a dedicated interview room and generally lasted 15-25 minutes. Patients were first asked to rate all potential attributes on a 0-10 scale, and then asked to identify any attributes that were missing from this list. Next, patients were asked to look at ten examples of risk graphics, each of which described the proportion of CLL patients that would respond to a hypothetical chemotherapy treatment. These ten examples included two examples of each icon in **Figure 1**, with one shaded sequentially and the other shaded at random. Colours were chosen to maximise contrast. Patients were asked to select the frequency statement that matched each graphic from four options. Some limited background data was also collected for each patient. All interview materials are provided in **Appendix 28**.

Figure 1: Potential graphical techniques for presenting risk information (clockwise from top-left: complex stick figures, simple stick figures, series of rectangles, grid of squares, grid of faces)



Results

15 patients were interviewed. No patients declined to be interviewed. Most patients answered all questions – exceptions are noted. No patients refused to answer questions. All patients appeared to understand the questions: the next question was not asked until the interviewer was confident that the previous question had been understood correctly. Clarifications and alternative explanations were provided if requested. **Table 1** summarises the characteristics of the patients that were interviewed.

Table 1: Patient characteristics (n=15 patients, except where indicated)

	Overall	Male	Female
Mean age (SD)	65.83 (10.60)	64.50 (13.23)	67.83 (5.16)
% male (n/N)	60% (9/15)	-	-
Mean school leaving age (SD)	17.14 (2.44)	17.63 (2.56) ^a	16.50 (2.35)
% employed (n/N)	20.00% (3/15)	22.22% (2/9)	16.67 (1/6)
Mean number of children (SD)	1.21 (1.31)	0.88 (1.25) ^a	1.67 (1.37)
Mean travel time to hospital in minutes (SD)	42.30 (29.43)	47.17 (26.36)	35.00 (34.75)
Mean time since diagnosis in years (SD)	6.47 (4.93)	6.78 (4.24)	6.00 (6.23)
% of patients currently receiving treatment (n/N)	13.33% (2/15)	11.11% (1/9)	16.67% (1/6)
Mean number of rounds of chemotherapy received across all patients (SD)	0.93 (0.80)	1.22 (0.83)	0.50 (0.55)
Mean courses of chemotherapy received across only those patients who have had chemotherapy (SD)	1.40 (0.78)	1.57 (0.53)	1.00 (0.00)
% of patients who have previously had a genetic test for CLL (n/N)	6.67% (1/15)	11.11% (1/9)	0.00% (0/6)

CLL = chronic lymphocytic leukaemia; SD = standard deviation. ^a One patient did not answer this question due to interviewer error

Patients were mainly male, aged 66, and retired. Most left school at the age of 17, and the average number of children per patient was 1.2. Most took around 45 minutes to reach hospital. In terms of clinical details, most patients were diagnosed with CLL six years ago. Only two patients were currently receiving treatment, but the average number of rounds of chemotherapy received across all patients was one. When patients who had never received chemotherapy were excluded, the average number of rounds of chemotherapy per patient rose to 1.4. Only one male patient recalled previously having a genetic test for CLL.

Attributes

Table 2 summarises the average attribute importance ratings and rankings, ordered by overall score. The highest rated attribute was ‘who explains the test result to you’. The lowest rated attribute was ‘number of times that a blood sample is required’. Answers were broadly similar by gender. The only attributes where there was a four-rank or greater difference between males and females were ‘test effectiveness’, and ‘level of information given to patients about the test’. In the former case, both males and females still ranked this attribute in the top five. No attributes outside of this longlist of nine attributes were identified by patients, although two made general comments about information on hereditary links, which are not relevant to this test. When asked to explicitly identify the most important parameter to them out of this longlist, most patients (8/12 – three non-responders) mentioned either ‘test effectiveness’, ‘test reliability’, or a combination of the two attributes. The next most common answer was ‘level of information’ (2/12). The value attached to the ‘location of test’ attribute was compared to the average travel time to hospital but results were similar for patients with high and low travel times.

Table 2: Average attribute scores and rankings (n=15 patients)

Attribute	Average importance rating ^a (ranking)		
	Overall (n=15)	Male (n=9)	Female (n=6)
Who explains the test result to you	8.2	8.6 (1)	7.7 (4)
Provision of exact diagnosis	8.1	8.0 (2)	8.3 (2)
Test reliability	7.5	7.0 (4)	8.2 (3)
Test effectiveness	7.2	6.6 (5)	8.4 (1)
Level of information given to patients about the test	6.4	7.2 (3)	5.2 (7)
Cost of test	6.2	6.1 (6)	6.3 (5)
Speed of test result	4.9	4.5 (8)	5.5 (6)
Location of test	4.7	5.1 (7)	4.2 (8)
Number of times that a blood sample is required	3.1	3.3 (9)	2.8 (9)

^a The rating scale was from 0 (not at all important) to 10 (extremely important)

Risk graphics

Table 3 reports the results of the risk graphic questions. Around two-thirds of all answers were correct, with males providing more correct answers than females. More correct answers were provided when graphics were ordered sequentially rather than non-sequentially, across both males and females. In terms of the most successful graphical format, the grid of faces returned the most correct answers across all patients. The other four graphical formats returned exactly the same results (around two-thirds correct) for males, so the overall results follow those reported for females. The second most successful format was therefore simple stick figures, followed by complex stick figures. Patients were also asked to state their most and least preferred graphical format. Several couldn't reach a decision about this. Of those who did answer (n=9), three expressed a preference for sequential graphics. There was no pattern to the remaining answers.

The number of correct answers was also compared to the average school leaving age. The average number of correct answers (out of ten) for patients who left school aged 18 or over (n=4) was 8.0. The average for patients who left school under the age of 18 (n=9) was 5.7.

Table 3: Risk interpretation (n=14 ^a)

Type of graphic	% of answers that were correct (absolute number)					
	Overall ^a	Ranking	Male	Ranking	Female	Ranking
All graphics	62% (87/140)	-	74% (59/80)	-	47% (28/60)	-
Sequential	74% (52/70)	-	85% (34/40)	-	60% (18/30)	-
Non-sequential	50% (35/70)	-	63% (25/40)	-	33% (10/30)	-
Complex stick figures	57% (16/28)	3	69% (11/16)	2=	42% (5/12)	3
Simple stick figures	61% (17/28)	2	69% (11/16)	2=	50% (6/12)	2
Grid of squares	54% (15/28)	4=	69% (11/16)	2=	33% (4/12)	4=
Series of rectangles	54% (15/28)	4=	69% (11/16)	2=	33% (4/12)	4=
Grid of faces	86% (24/28)	1	94% (15/16)	1	75% (9/12)	1

^a One patient did not provide responses in this part of the interview

Discussion

Interviewees were typical of CLL patients in general: predominantly male, over 60 years of age, diagnosed for over six years and mostly not undergoing treatment. Although the number of patients interviewed was small (hence no tests of statistical significance are performed), these

patients do represent around 10% of all CLL patients seen within the OUH NHS Trust, hence these results were cautiously taken to be representative of the patient group as a whole. In addition, the literature suggests that saturation in such interviews is commonly reached at very low sample sizes [358].

Attributes

Based on these results, and taking into account how many attributes would be manageable for this patient population, how many would be required to enable informed choices, and which could best capture the characteristics of current and future testing practice, six attributes were selected for the final DCE. 'Test reliability' and 'test effectiveness' were ranked highly by patients and explicitly identified as being important factors in their decision to undergo genetic testing, so were retained. 'Cost of test' is a crucial attribute as it enables the calculation of willingness-to-pay, hence this was also retained for the final DCE. 'Speed of test result' was one of the lower ranked variables, however, this was also retained for the final DCE as it provides an alternative continuous variable (other than cost) which could be used to calculate marginal rates of substitution. 'Who explains the test result to you' was ranked highest across all patients, so this was the fifth attribute retained for the final DCE. Finally, 'level of information given to patients about the test' was ranked fifth overall and captured important information related to test complexity (and how much of this complexity patients wished to understand), so was also retained.

Three attributes were not retained. 'Provision of exact diagnosis' was included on the recommendation of clinicians and was ranked highly by patients. However, the provision of a more exact diagnosis was ultimately not considered to be a primary objective for genetic and genomic testing in this context, so this attribute was dropped. 'Location of test' was dropped as it was not valued highly by patients. Those who did rank this attribute highly found it difficult to separate location from who provided the test result, judging by comments made during the interview. Finally, 'number of times that a blood sample is required' was the lowest valued attribute for both males and females.

Risk graphics

The results for the graphical format used to display risk information were more clear-cut. Sequential graphics returned more correct answers than non-sequential graphics, across all patients. Of the different types of icon, a grid of faces returned more correct answers than any of the other options. A grid of faces was therefore chosen as the comparator in the final DCE, with information presented sequentially instead of non-sequentially.

Appendix 30: Further information on the pilot discrete choice experiment and the final experimental design

The pilot discrete choice experiment (DCE) aimed to model the probability that a respondent would select a specified genomic test. The probability of choosing a given alternative was determined by the indirect utility function (V):

$$V = \alpha.ASC + \beta_1.TIME + \beta_2.COST + \beta_3.EFFECT + \beta_4.REL + \beta_5.INFO0 + \beta_6.INFO1 + \beta_7.INFO2 + \beta_8.WHO0 + \beta_9.WHO1 + \beta_{10}.WHO2 + \beta_{11}.COST.EFFECT + \beta_{12}.COST.REL + \gamma.Z + \varepsilon$$

Where: α = estimate for the alternative specific constant (ASC) for the first choice (not expected to be significant, as the choices are unlabelled)

β_1 to β_{12} = parameters

Z = set of covariate terms of interest

γ = set of estimates for these covariate terms

ε = error term (unobservable random component)

A fractional factorial design was produced using Ngene [276]. This design minimised the d-error for a multinomial logit model (MNL) with no dominated alternatives, minimal overlap, and perfect level balance. The MNL regression results are presented below:

Coefficient	Variable	Coefficient value	Standard error
α	ASC	0.13505	0.18396
β_1	TIME	0.00843	0.02686
β_2	COST	-0.00039	0.00169
β_3	EFFECT	0.02958*	0.00963
β_4	REL	-0.11290	0.09906
β_5	INFO0 (5 minutes)	-0.25321	0.20344
β_6	INFO1 (10 minutes)	0.08554	0.19181
β_7	INFO2 (15 minutes)	0.35424	0.20328
-	INFO3 (20 minutes) ^a	-0.18657*	0.00317
β_8	WHO0 (general practitioner)	-0.51833*	0.19214
β_9	WHO1 (specialist nurse)	0.04017	0.20312
β_{10}	WHO2 (junior hospital doctor)	-0.09470	0.18661
-	WHO3 (consultant hospital doctor) ^a	0.57287*	0.00974
β_{11}	COST x EFFECT	-0.00002	0.00002
β_{12}	COST x REL	-0.00005	0.00025

Pseudo r-squared = 0.2188. * P-value < 0.01. ^a Coefficients calculated as the negative sum of the coefficients of the other three levels; standard errors calculated as the average of the other three coefficients for this attribute.

Coefficient signs were mostly as expected, with negative effects on COST and REL and a significant positive effect on EFFECT. The TIME attribute had a small positive coefficient (possibly counter-intuitive) but was not significant. Both interaction terms had small negative coefficients and were not significant. Patients expressed an increasing preference for more information which peaked at 15 minutes. Of the different types of clinician who could explain the test result to patients, a consultant hospital doctor had the largest positive coefficient, while a general practitioner had the largest negative coefficient.

The final experimental design was again produced using Ngene and incorporated the results of the pilot regression as priors. This final design took into consideration the fact that respondent heterogeneity might be important in this context: it is likely that there will be considerable variation

in attitudes towards risk amongst respondents, for example. As one of the objectives of the DCE was to evaluate the impact of presenting risk information in visual and non-visual formats, it was considered to be important to generate an experimental design which permits the use of more flexible regression models to evaluate respondent choices (such as the mixed logit approach, which can capture respondent heterogeneity). For this reason, a model averaging approach was employed for this final experimental design which compared the d-error of five models, all weighted equally:

1. MNL with no priors;
2. MNL with priors for significant variables only;
3. MNL with priors for all variables;
4. Mixed multinomial logit (MMNL) with significant variables (EFFECT and WHO0) defined as random draws from normal distributions and the remaining variables defined with zero priors;
5. MMNL with significant variables (EFFECT and WHO0) defined as random draws from normal distributions and the remaining variables defined by mean values (no random draw).

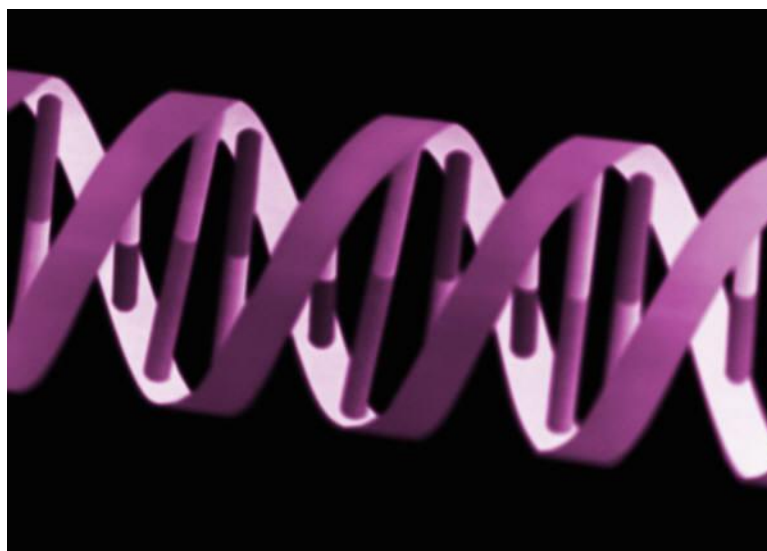
The random draws were undertaken using 500 Halton draws, which returned a lower d-error more quickly than other approaches. Panel MMNL models were also considered, with either only significant or all parameters defined as random draws, as well as MMNL models with all parameters defined as random draws. However, Ngene was unable to process these models due to the complexity of the DCE design.

The final specification again minimised the d-error and had no dominated alternatives, minimal overlap, and perfect level balance. Given that priors now informed the experimental design, the choice probabilities for the chosen design were also considered. When no prior information is available, all choices are effectively 50:50 and respondents will be perfectly indifferent between the two alternatives (and therefore more uncertain about their responses). Once prior information is available on likely strength of preference for different attribute levels, choice probabilities can vary. Probabilities close to 50:50 or 100:0 are least informative: these choices are likely to either be very easy for respondents (100:0), or very difficult (50:50). Choices with probabilities close to 75:25 (i.e. of moderate difficulty) are most informative. The proportion of choice probabilities for the final design that were either close to 50:50 (within 55:45 either way) or close to 100:0 (95:5 either way) was calculated to determine how informative the final design was likely to be. This proportion was 17%, which was determined to be acceptable.

Appendix 31: Discrete choice experiment survey with risk graphics

Genomic testing and chronic lymphocytic leukaemia (CLL)

Do patients value tests which can predict their response to chemoimmunotherapy?



Doctors and researchers in Oxford are carrying out a survey about a new test which could be offered to chronic lymphocytic leukaemia (CLL) patients. We are interested in your views about how this new test could be provided by the NHS. Your participation is entirely voluntary, but your views would help us to develop this testing service in the future.

Please try to answer all of the questions in this booklet. There are no right answers - we are just interested in your views.

This questionnaire takes around 20 minutes to complete. Please return your completed questionnaire to James Buchanan, or, if he is not around, please hand it in to the reception staff. If you do not have time to complete the questionnaire today, please complete it at home and send it to us by post free of charge, using the stamped addressed envelope that we have provided.

If you have any questions about this questionnaire, please contact:

James Buchanan

NIHR Doctoral Research Fellow

Health Economics Research Centre, Department of Public Health
University of Oxford, Old Road Campus

Headington

Oxford

OX3 7LF

Telephone: 01865 289262 Email: oxfordcllsurvey@dph.ox.ac.uk

The doctor who is leading this project is **Dr Anna Schuh** (MD, PhD, MRCP, FRCPATH) Consultant Haematologist, Oxford Cancer and Haematology Centre, Churchill Hospital, Headington, Oxford, OX3 7LE (Telephone: 01865 235259, Email: anna.schuh@nhs.net).

For more information about this study, visit: www.herc.ox.ac.uk/ctl

HOW TO COMPLETE THIS QUESTIONNAIRE

We would like you to imagine that you are a CLL patient who requires chemoimmunotherapy treatment (e.g. combined fludarabine, cyclophosphamide, and rituximab (FCR) treatment). Out of every **400** CLL patients, we expect that about **300** will respond really well to the treatment that is usually tried first, and **100** will not respond (or will relapse within 2 years of treatment). **There is no way of knowing which group you are in before you start treatment.**

If we could identify the patients that will not respond to the usual chemoimmunotherapy treatment before this treatment is started, we might be able to avoid giving them treatment that will not work: an alternative treatment would be offered instead.

Next, we would also like you to imagine that you have been offered a test which could provide you and your doctors with some information about how you will respond to the usual treatment. This test is called a **genomic test.**

This genomic test works by identifying variations in the genetic make-up (DNA) of your leukaemia cells. These variations affect how you will respond to the usual chemoimmunotherapy treatment. This information helps your doctors to decide whether to give you the **usual chemoimmunotherapy treatment** or whether to give you an **alternative chemoimmunotherapy treatment.** This test **does not** look at inheritable genetic information, it only looks at the genes in the leukaemia cells.

In this survey you will be presented with several **imaginary** situations in which two different tests will be described. In each situation, we would like you to choose which test you prefer out of two options: **Test A** or **Test B.** Even if you do not like either test, please mark the one you would prefer if you had to choose one.

Each imaginary test will be described to you in terms of six features:

- Time to receive the test result
- Cost of the test
- Ability of the test to predict who will not respond to the usual chemotherapy treatment
- Test reliability
- Length of time health care professionals spend describing the test and the test results to you
- Type of health care professional who explains the test result to you

A detailed description of each feature is on the next page. Please read these descriptions carefully, then try to answer all of the questions that follow.

There are no right or wrong answers - we are just interested in your views. All information will be kept strictly confidential and used for research purposes only.

THANK YOU FOR YOUR TIME AND ASSISTANCE

FEATURES

Each imaginary test will be described to you in terms of the following features:

- **Time to receive test result**

This describes how many days are between you providing a blood sample and receiving the test result. You will only begin a chemoimmunotherapy treatment once you receive the test result.

- **Cost of test to you**

This describes how much it costs for one test. Obviously in the NHS you are not asked to pay for medical tests – the NHS covers this cost. However, for the purposes of this survey, imagine that you live in a country where you have to cover the cost of your medical tests out of your own pocket, such as the United States of America (USA).

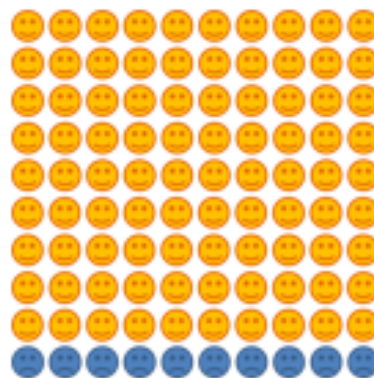
- **Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment**

Out of 400 CLL patients like you, we expect that about **300** will respond really well to the chemoimmunotherapy treatment that is usually tried first, and **100** will not respond (or will relapse within 2 years of treatment). Some genomic tests can detect most of the 100 patients who will not respond. Other genomic tests can only detect a few of these patients. The ability of the genomic test to predict if you are one of the patients who will not respond to usual treatment is given by this feature.

A figure of 30 means that the genomic test will pick up 30 out of every 100 CLL patients who will not respond to usual treatment.

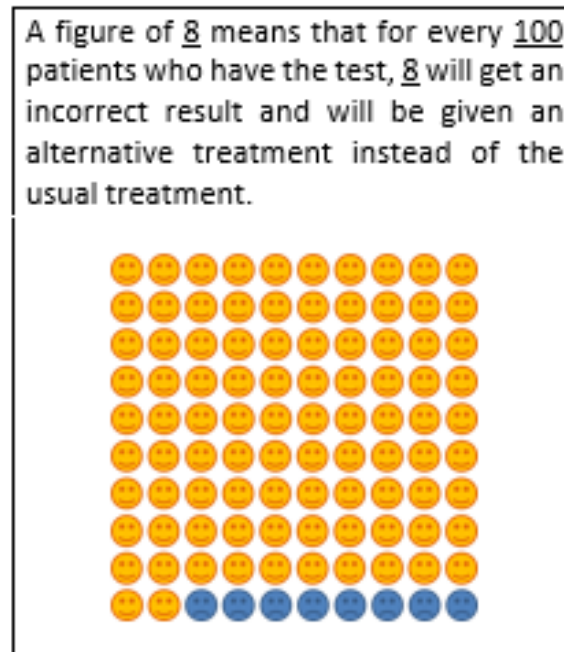
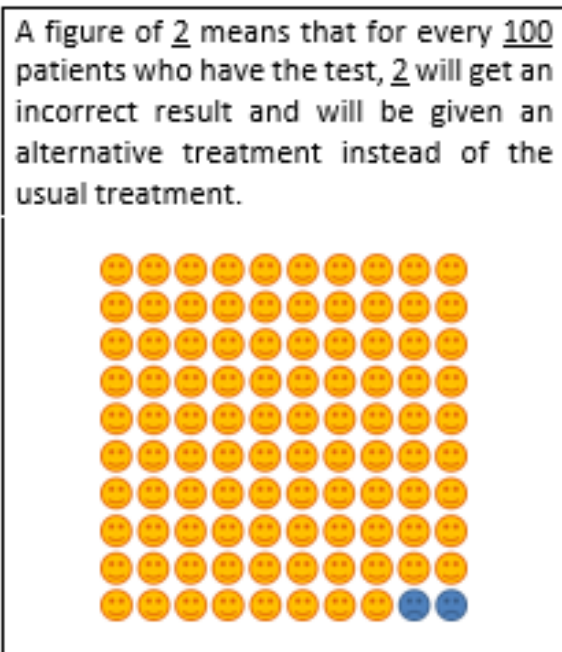


A figure of 90 means that the genomic test will pick up 90 out of every 100 patients who will not respond to usual treatment.



- **Test reliability**

Occasionally a genomic test will provide an incorrect result: the test result suggests that you won't respond to the usual chemoimmunotherapy treatment when in fact you will. This means that you will incorrectly receive an alternative treatment instead of the usual treatment. This alternative treatment may be slightly less effective than the usual treatment. The most reliable tests rarely provide an incorrect result. Less reliable tests return incorrect results more often. The reliability of the genomic test is given by this feature.



- **Length of time health care professionals spend describing the test and the test results to you**

5 minutes: Before the test, your health care professional tells you that they will be taking a blood sample in order to do a genomic test that will provide information on whether you will respond to the usual chemoimmunotherapy treatment. After you have had the test, he/she tells you whether the test has indicated that you will respond to the usual treatment, and recommends an alternative treatment if necessary.

10 minutes: In addition to the information provided in 5 minutes, your health care professional also provides some limited information about how the test works and the different chemoimmunotherapy treatments.

15 minutes: In addition to the information provided in 10 minutes, your health care professional also describes in detail the consequences of different chemoimmunotherapy treatments in terms of side-effects, success rates and prognosis.

20 minutes: In addition to the information provided in 15 minutes, your health care professional also provides detailed test results (explaining which genetic mutations have been found, if any) and has time to answer any questions you may have about any potential consequences for your family.

- **Type of health care professional who explains the test result to you**

General practitioner: This is the GP who you would normally go to see in your local doctors surgery or health centre.

Specialist nurse: This is a nurse who specialises in CLL and who works in the outpatient clinic of the hospital that you normally go to for check-ups and treatments.

Junior hospital doctor: This is a doctor who works in the outpatient clinic of the hospital that you normally go to for check-ups and treatments. He or she has some knowledge of CLL.

Consultant hospital doctor: This is a doctor who works in the outpatient clinic of the hospital that you normally go to for check-ups and treatments. He or she is an experienced CLL specialist.

DESCRIBING A GENOMIC TEST

Each imaginary test will always be described to you in terms of the same six features. Each feature has four possible levels. These levels will change from one question to the next. Please read through the descriptions of the features and levels.

Feature	Possible levels
Time to receive test result	5 days
	8 days
	11 days
	14 days

Cost of test to you (remember, in the NHS you would not have to actually pay for these tests)	£130
	£260
	£400
	£600

Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment
	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment
	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment
	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment

Test reliability (number of incorrect results provided every 100 tests)	<u>2</u> out of every <u>100</u> tests provide an incorrect result
	<u>4</u> out of every <u>100</u> tests provide an incorrect result
	<u>6</u> out of every <u>100</u> tests provide an incorrect result
	<u>8</u> out of every <u>100</u> tests provide an incorrect result

Length of time health care professionals spend describing the test and the test results to you	5 minutes
	10 minutes
	15 minutes
	20 minutes

Type of health care professional who explains the test result to you	General practitioner
	Specialist nurse
	Junior hospital doctor
	Consultant hospital doctor

Once you have read through the descriptions, we would like to know which of these features are most important to you. Please complete **QUESTION 1** on the next page.

QUESTION 1: FEATURE RANKING EXERCISE

Please consider all 6 of the features that make up the test and rank these in order of importance to you in the yellow boxes.

Rank the feature that is most important to you with a 1. Then rank the second most important feature with a 2. Do the same for the third, fourth and fifth most important features. Rank the feature that is least important to you with a 6.

Feature	Rank
Time to receive the test result	
Cost of the test <u>to you</u>	
Ability of the test to predict who will <u>not</u> respond to the usual chemoimmunotherapy treatment	
Test reliability (number of incorrect results provided every 100 tests)	
Length of time health care professionals spend describing the test and the test results to you	
Type of health care professional who explains the test result to you	

Once you have ranked these features, read through the **EXAMPLE QUESTION** on the next page. This will provide some guidance on answering the questions.

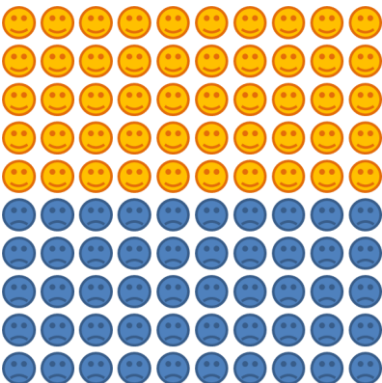
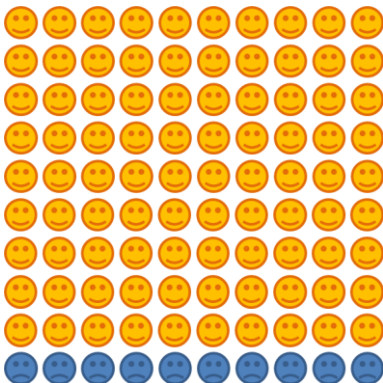
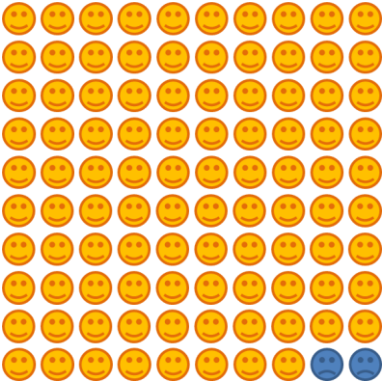
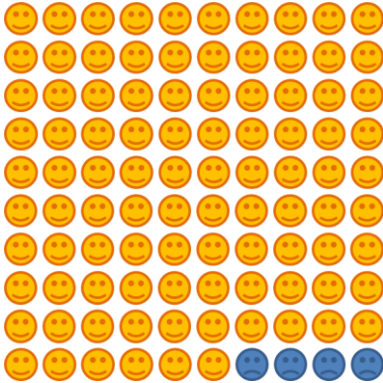
Then go to **QUESTION 2**.

Indicate with a cross (X) whether you would prefer **TEST A** or **TEST B**. You can do this by writing X in the relevant yellow box.

Repeat this for **QUESTIONS 3 to 18**

EXAMPLE QUESTION – DO NOT ANSWER

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	5 days
Cost of the test to you	£260	£600
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result 	<u>4</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	10 minutes	5 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	Junior hospital doctor

Which test would you prefer?

☐

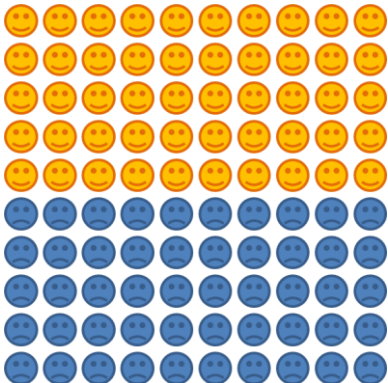
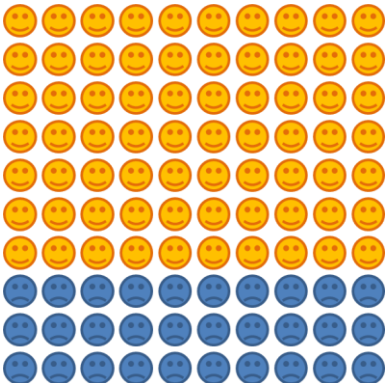
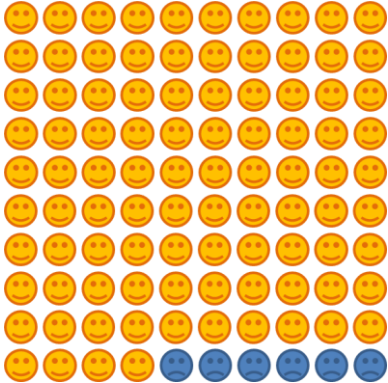
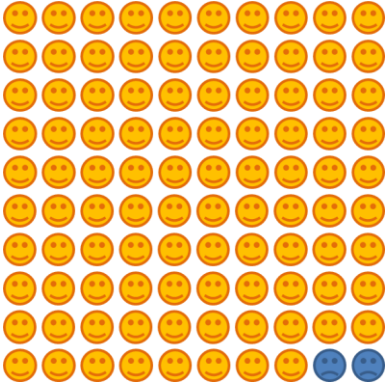
TEST A

☒

TEST B

QUESTION 2

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	5 days
Cost of the test to you	£260	£130
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>6</u> out of every <u>100</u> tests provide an incorrect result 	<u>2</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	10 minutes	15 minutes
Type of health care professional who explains the test result to you	General practitioner	Consultant hospital doctor

Which test would you prefer?



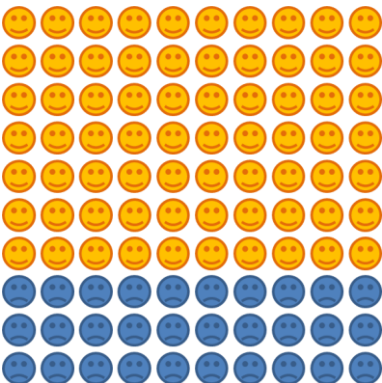
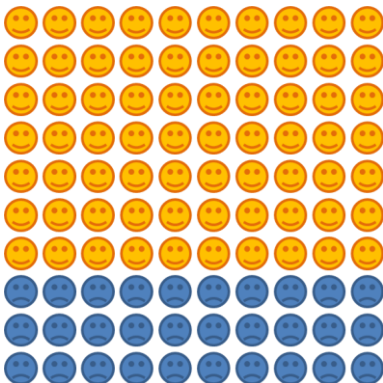
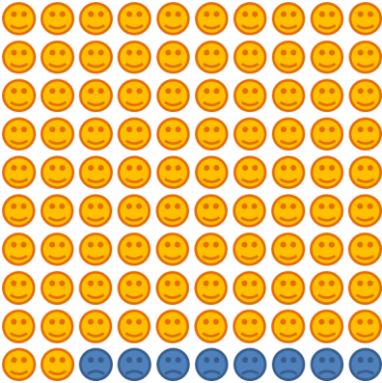
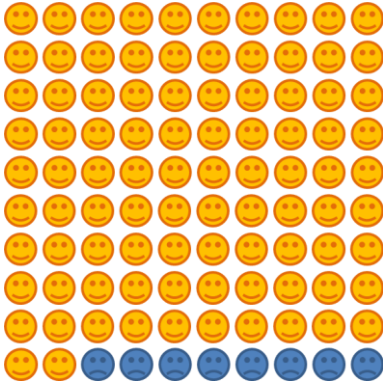
TEST A



TEST B

QUESTION 3

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	5 days	14 days
Cost of the test to you	£130	£600
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>8</u> out of every <u>100</u> tests provide an incorrect result 	<u>8</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	5 minutes	10 minutes
Type of health care professional who explains the test result to you	Specialist nurse	Junior hospital doctor

Which test would you prefer?



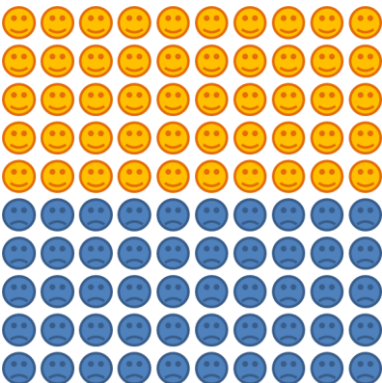
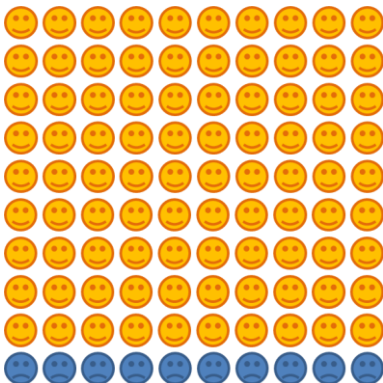
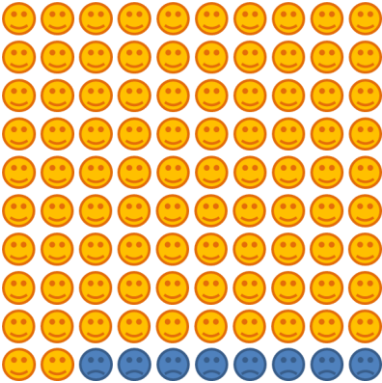
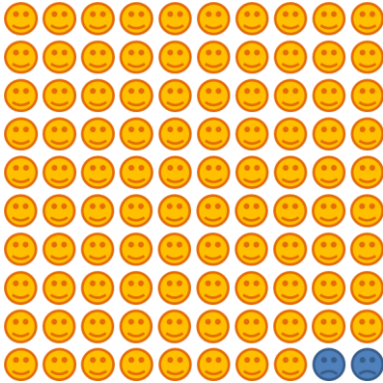
TEST A



TEST B

QUESTION 4

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	5 days	14 days
Cost of the test to you	£400	£130
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>8</u> out of every <u>100</u> tests provide an incorrect result 	<u>2</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	10 minutes	15 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	General practitioner

Which test would you prefer?



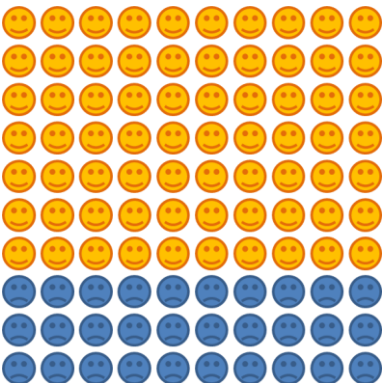
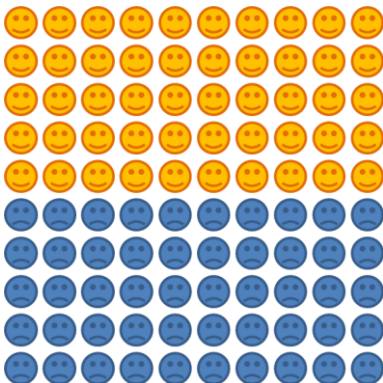
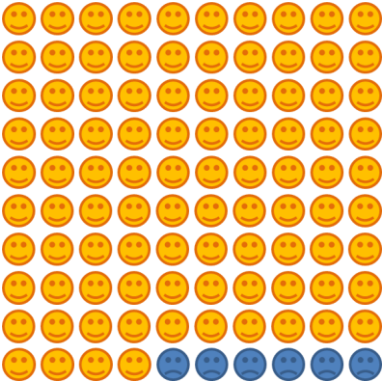
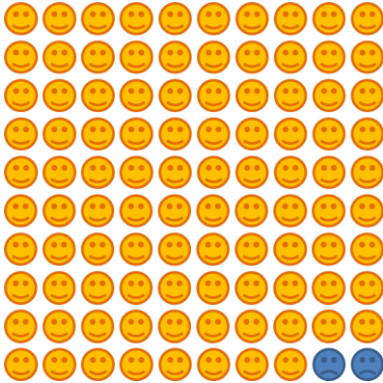
TEST A



TEST B

QUESTION 5

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	14 days	5 days
Cost of the test to you	£260	£400
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>6</u> out of every <u>100</u> tests provide an incorrect result 	<u>2</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	10 minutes	20 minutes
Type of health care professional who explains the test result to you	Specialist nurse	Junior hospital doctor

Which test would you prefer?



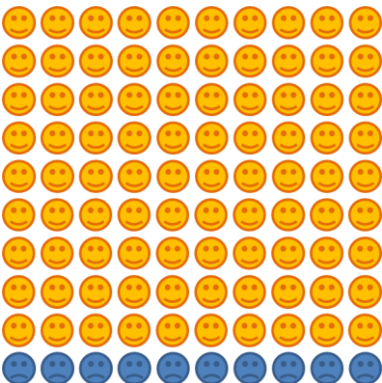
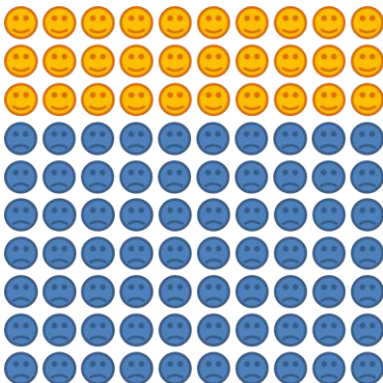
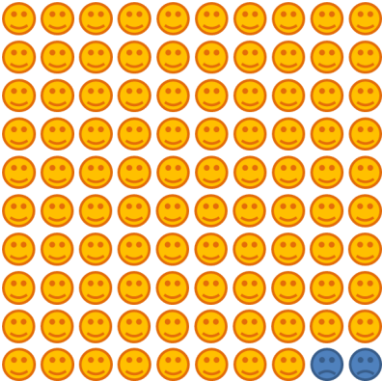
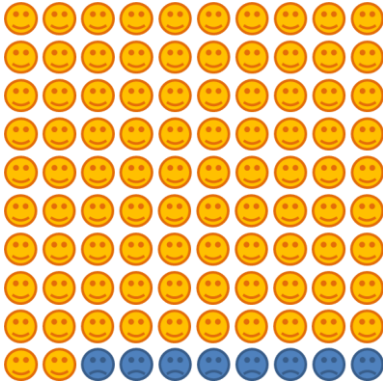
TEST A



TEST B

QUESTION 6

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	11 days
Cost of the test to you	£600	£600
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result 	<u>8</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	10 minutes	20 minutes
Type of health care professional who explains the test result to you	General practitioner	Specialist nurse

Which test would you prefer?



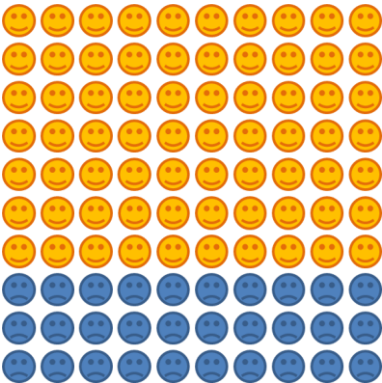
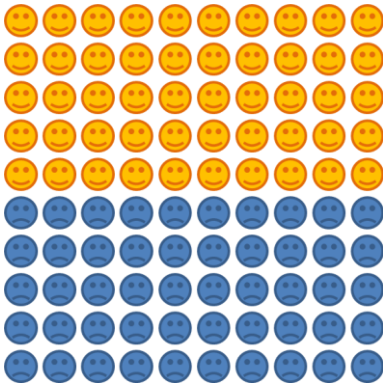
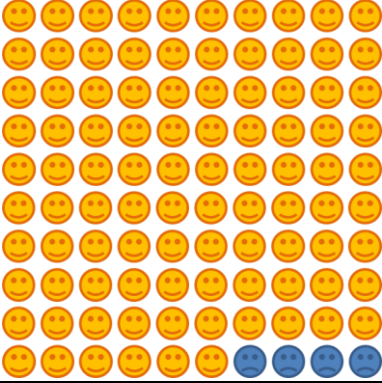
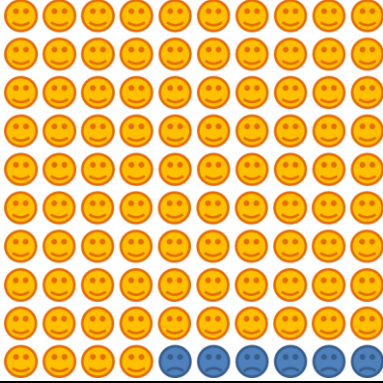
TEST A



TEST B

QUESTION 7

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

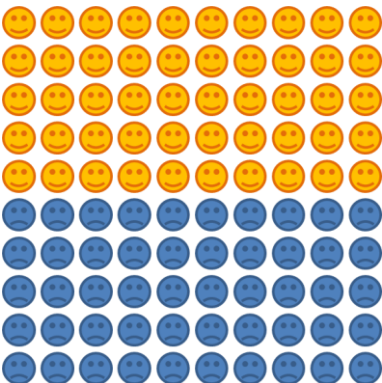
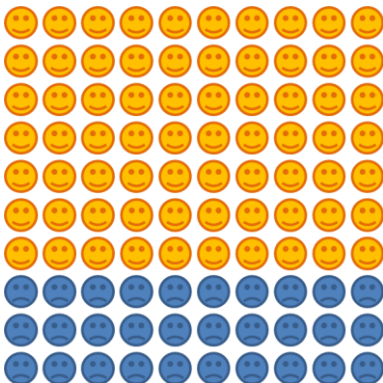
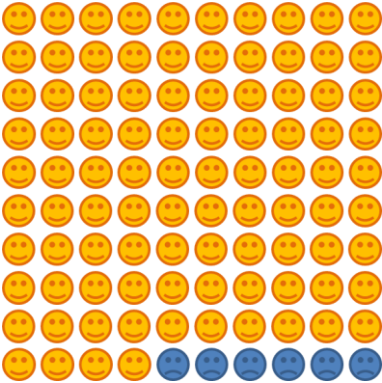
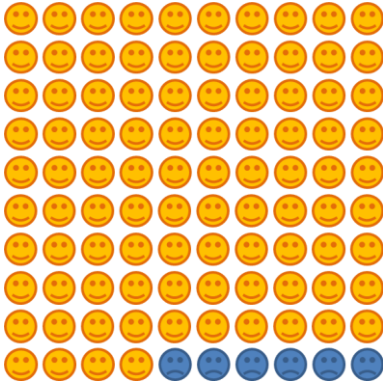
Feature	Test A	Test B
Time to receive the test result	5 days	14 days
Cost of the test to you	£400	£260
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>4</u> out of every <u>100</u> tests provide an incorrect result 	<u>6</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	20 minutes	5 minutes
Type of health care professional who explains the test result to you	Specialist nurse	General practitioner

Which test would you prefer?



QUESTION 8

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	14 days	5 days
Cost of the test to you	£260	£400
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>6</u> out of every <u>100</u> tests provide an incorrect result 	<u>6</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	20 minutes	15 minutes
Type of health care professional who explains the test result to you	Junior hospital doctor	Consultant hospital doctor

Which test would you prefer?



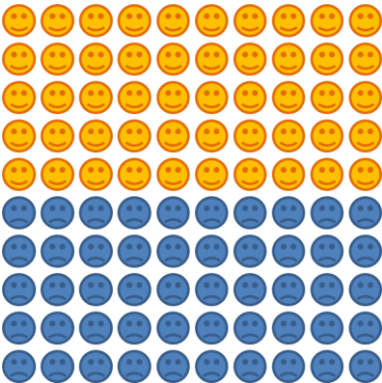
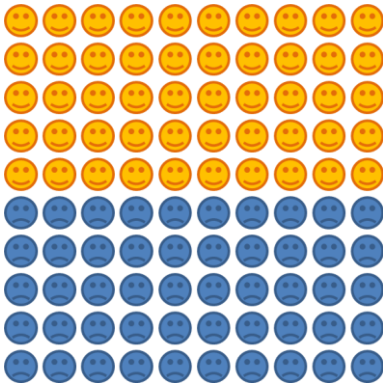
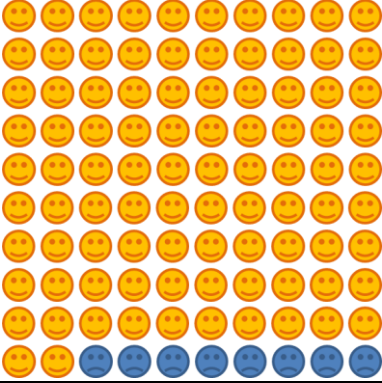
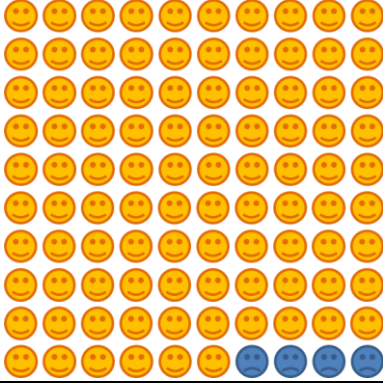
TEST A



TEST B

QUESTION 9

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	14 days	5 days
Cost of the test to you	£130	£400
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>8</u> out of every <u>100</u> tests provide an incorrect result 	<u>4</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	20 minutes	5 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	Specialist nurse

Which test would you prefer?



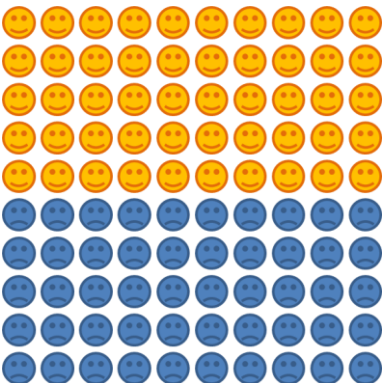
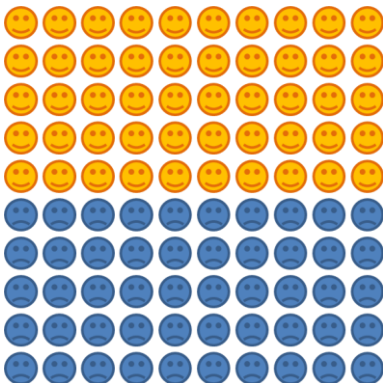
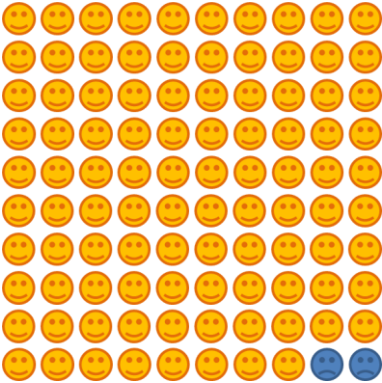
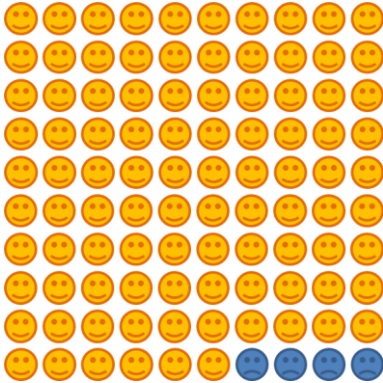
TEST A



TEST B

QUESTION 10

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	14 days	5 days
Cost of the test to you	£600	£260
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result 	<u>4</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	5 minutes	20 minutes
Type of health care professional who explains the test result to you	Junior hospital doctor	General practitioner

Which test would you prefer?



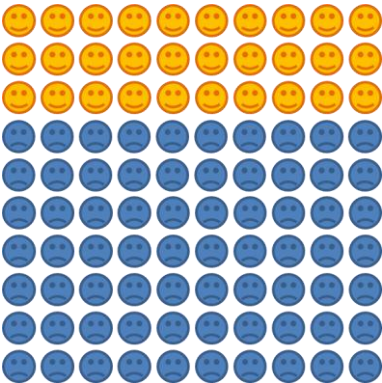
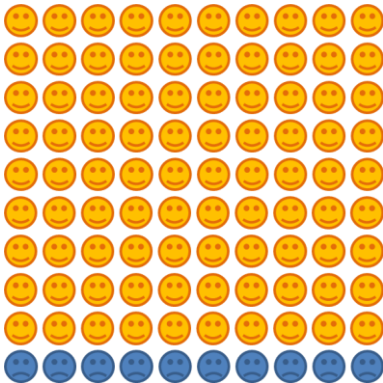
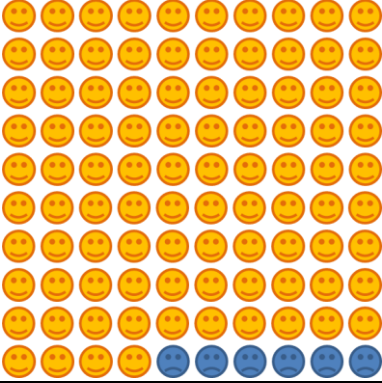
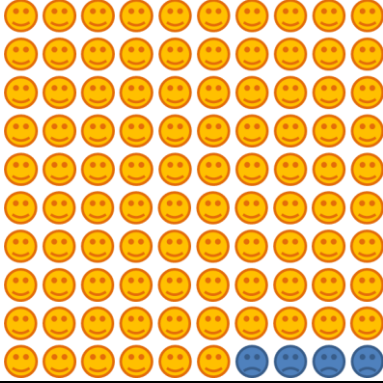
TEST A



TEST B

QUESTION 11

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	5 days	14 days
Cost of the test to you	£400	£600
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>6</u> out of every <u>100</u> tests provide an incorrect result 	<u>4</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	15 minutes	20 minutes
Type of health care professional who explains the test result to you	Junior hospital doctor	Specialist nurse

Which test would you prefer?



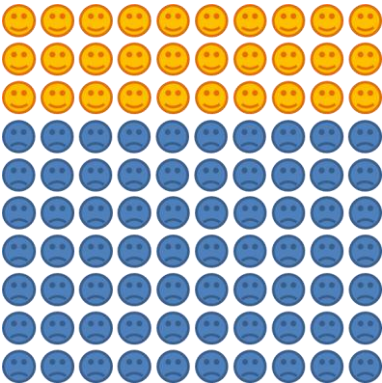
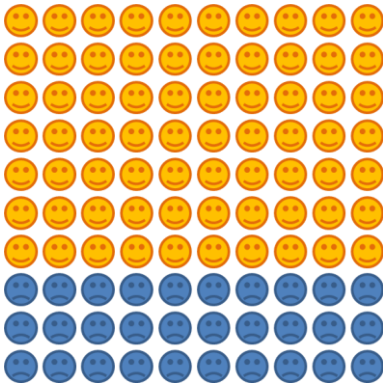
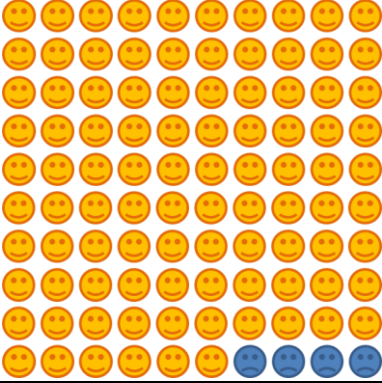
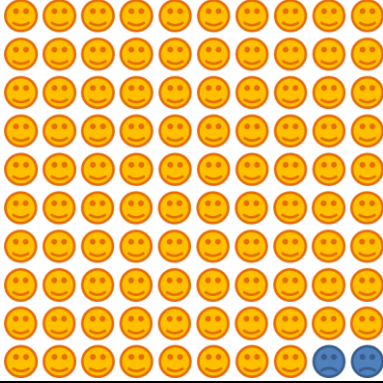
TEST A



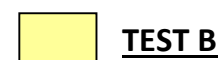
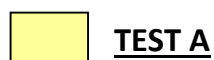
TEST B

QUESTION 12

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

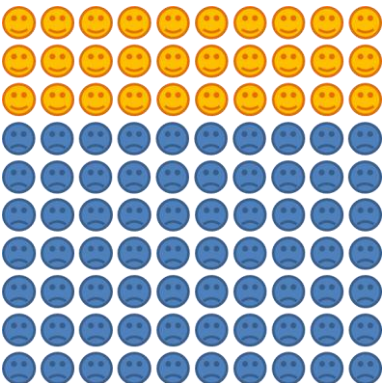
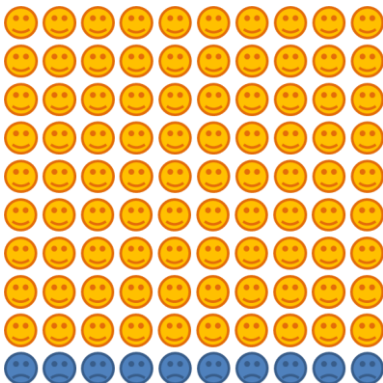
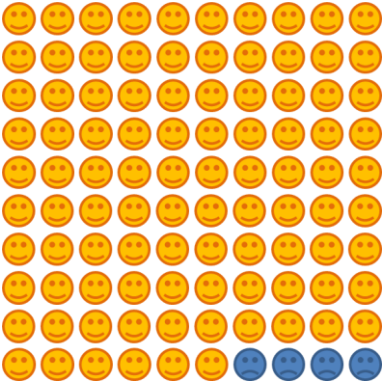
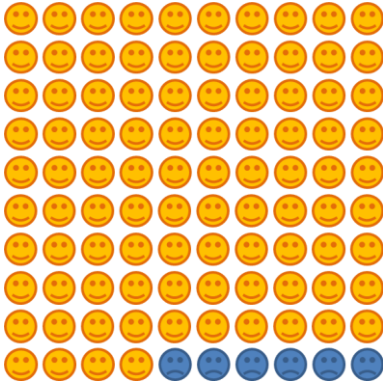
Feature	Test A	Test B
Time to receive the test result	11 days	8 days
Cost of the test to you	£600	£130
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>4</u> out of every <u>100</u> tests provide an incorrect result 	<u>2</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	15 minutes	10 minutes
Type of health care professional who explains the test result to you	General practitioner	Junior hospital doctor

Which test would you prefer?



QUESTION 13

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	11 days	8 days
Cost of the test to you	£260	£400
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>4</u> out of every <u>100</u> tests provide an incorrect result 	<u>6</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	10 minutes	5 minutes
Type of health care professional who explains the test result to you	Specialist nurse	Consultant hospital doctor

Which test would you prefer?



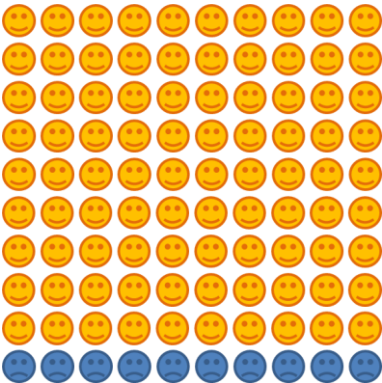
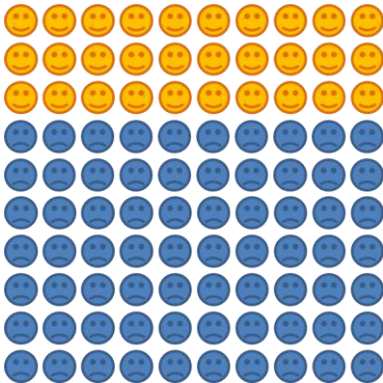
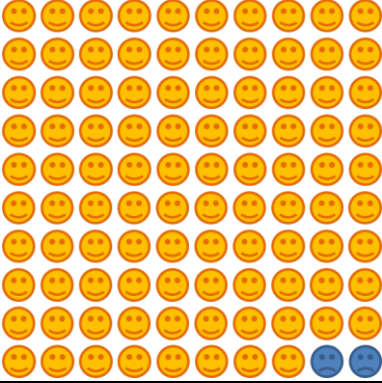
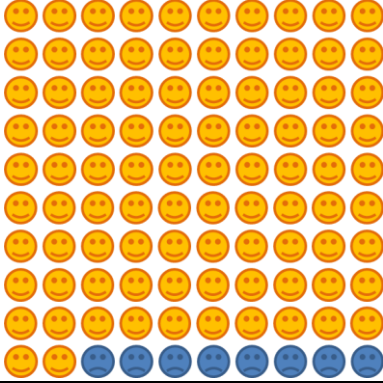
TEST A



TEST B

QUESTION 14

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	11 days
Cost of the test to you	£130	£130
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result 	<u>8</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	5 minutes	15 minutes
Type of health care professional who explains the test result to you	General practitioner	Specialist nurse

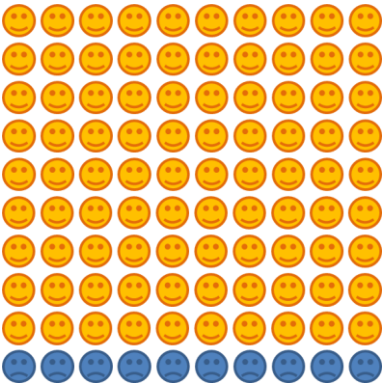
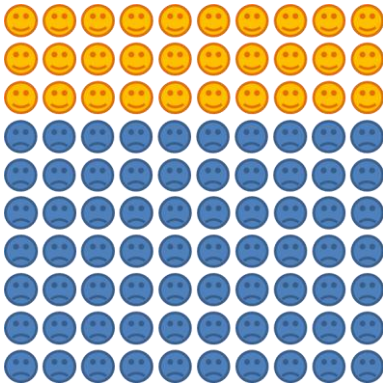
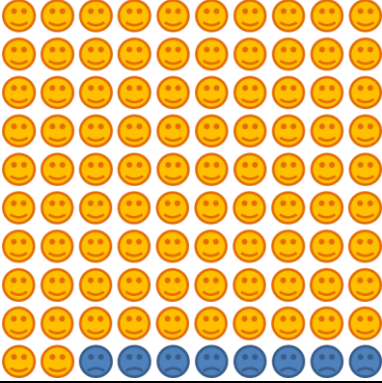
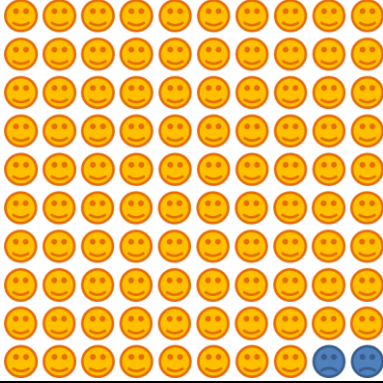
Which test would you prefer?

☐ **TEST A**

☐ **TEST B**

QUESTION 15

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	11 days
Cost of the test to you	£600	£600
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>8</u> out of every <u>100</u> tests provide an incorrect result 	<u>2</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	15 minutes	10 minutes
Type of health care professional who explains the test result to you	Junior hospital doctor	Consultant hospital doctor

Which test would you prefer?



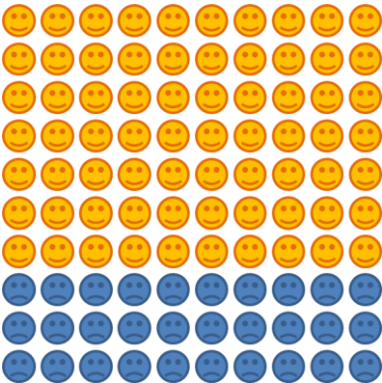
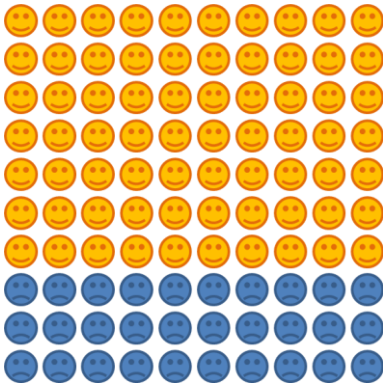
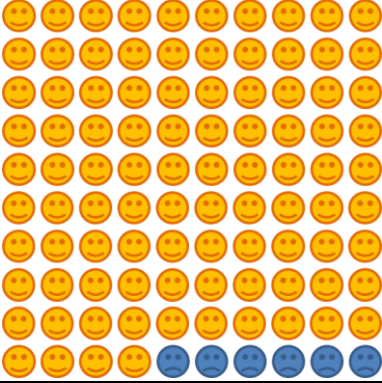
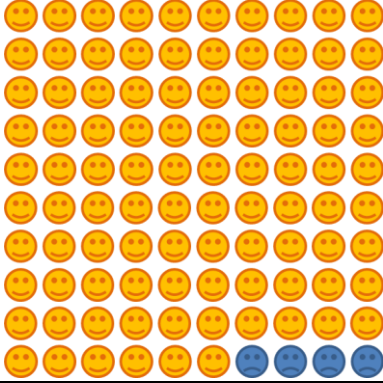
TEST A



TEST B

QUESTION 16

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	11 days
Cost of the test to you	£400	£260
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>6</u> out of every <u>100</u> tests provide an incorrect result 	<u>4</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	20 minutes	15 minutes
Type of health care professional who explains the test result to you	General practitioner	Consultant hospital doctor

Which test would you prefer?



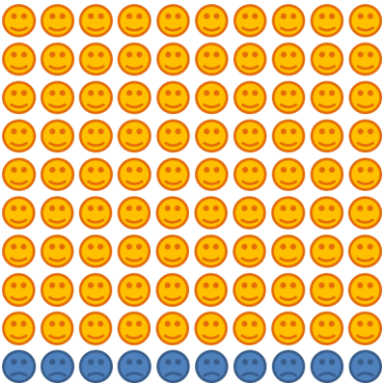
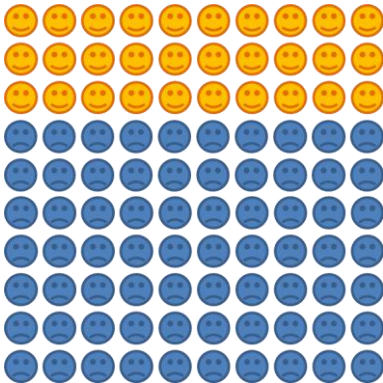
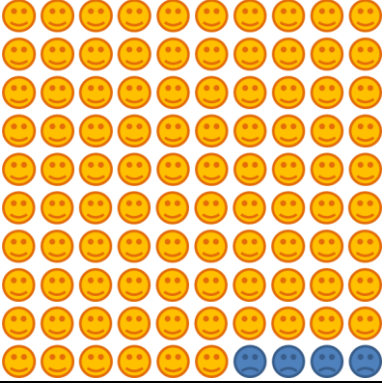
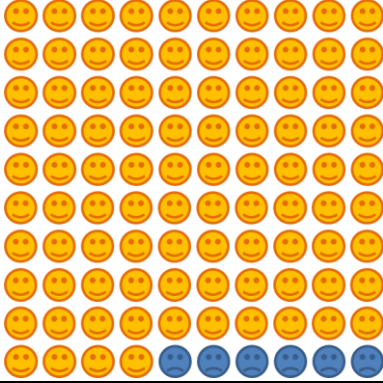
TEST A



TEST B

QUESTION 17

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	11 days	8 days
Cost of the test to you	£260	£260
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>4</u> out of every <u>100</u> tests provide an incorrect result 	<u>6</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	15 minutes	5 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	General practitioner

Which test would you prefer?



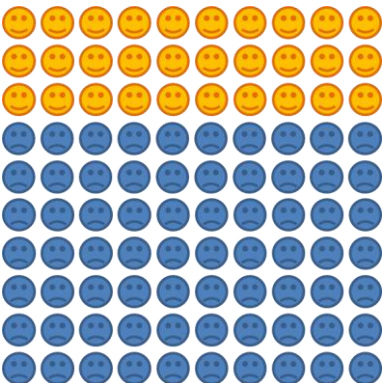
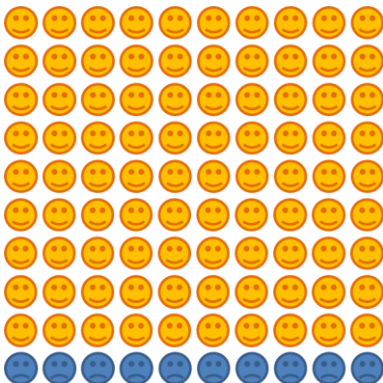
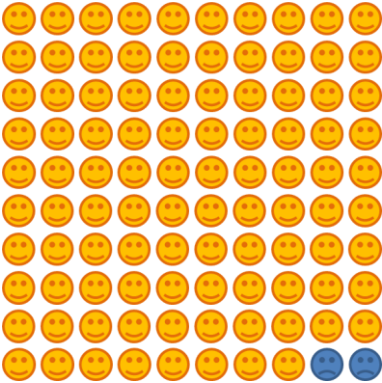
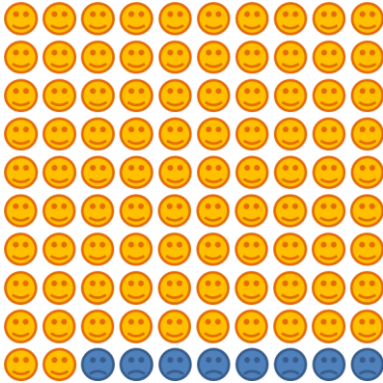
TEST A



TEST B

QUESTION 18

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	11 days	8 days
Cost of the test to you	£130	£130
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment 	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment 
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result 	<u>8</u> out of every <u>100</u> tests provide an incorrect result 
Length of time health care professionals spend describing the test and the test results to you	5 minutes	10 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	Junior hospital doctor

Which test would you prefer?



TEST A



TEST B

QUESTION 19: BACKGROUND INFORMATION

We would now like to ask some questions about **you**. All the information that you provide will help us in our analysis, and all of your details will be kept confidential. If you do not wish to answer some of these questions you do not have to.

a) What is your gender? ☐ Male ☐ Female

b) What is your age? years old

c) What is your occupational status?

☐ Employed full time

☐ Employed part-time

☐ Self-employed

☐ Unemployed

☐ Retired

☐ Homemaker/Stay at home parent

☐ Student

☐ Other

d) How old were you when you left full time education? years old

e) Please estimate the annual income of your household before tax and national insurance:

☐ Less than £10,000

☐ £10,000 - £19,999

☐ £20,000 - £29,999

☐ £30,000 - £39,999

☐ £40,000 and over

f) How many people live in your household? people

g) How many children do you have? Living at home Living away from home

h) How long does it take you to reach the hospital that you usually attend for treatment?

☐ 0-15 minutes

☐ 16-30 minutes

☐ 31-45 minutes

☐ 46-60 minutes

☐ 61-75 minutes

☐ 76-90 minutes

☐ More than 90 minutes

i) What mode of transport do you normally take to reach the hospital that you usually attend for treatment? Please select the mode on which you complete most of your journey.

☐ Hospital arranged transport (e.g. ambulance, hospital arranged taxi)

☐ Public transport (e.g. privately arranged taxi, bus, train)

☐ Private transport (e.g. private car, motorcycle, bicycle, walk)

j) How long ago were you diagnosed with CLL, to the nearest year? years

k) Have you ever undergone chemoimmunotherapy treatment for your CLL? ☐ Yes (*go to question 1*)

☐ No (*go to question 20*)

l) If yes, how many times have you undergone a course of chemoimmunotherapy treatment? If you have received one type of chemoimmunotherapy several times over a period of months, this counts as one course of treatment. times

m) Have you ever stayed in hospital overnight as a result of undergoing chemoimmunotherapy treatment? ☐ Yes ☐ No

n) Are you currently undergoing chemoimmunotherapy treatment for your CLL? ☐ Yes (go to question o) ☐ No (go to question 20)

o) If yes, what type of treatment are you currently receiving?

QUESTION 20: GENETIC AND GENOMIC TESTING

a) Do you know if you have had a genetic or genomic test since being diagnosed with CLL (sometimes these are called FISH tests, or tests to see if you have a certain gene mutation)? ☐ Yes (go to question b) ☐ No (go to question d) ☐ Not sure (go to question d)

b) If yes, do you know what type of genetic test this was?

c) Did you find the information provided by this test to be useful?

☐ Yes ☐ No ☐ Not sure

d) Do you know anybody (e.g. a family member or a friend) who has had a genetic or genomic test? ☐ Yes ☐ No

e) Please indicate whether you agree or disagree with each of these statements. Place a cross (X) by either Yes, No or Not sure for each statement.

Statement	Yes	No	Not sure
I am in favour of using genetic or genomic testing to identify CLL patients who will not respond to usual chemoimmunotherapy treatment	☐	☐	☐
I am in favour of using genetic or genomic testing to identify patients with other diseases who will not respond to usual treatments	☐	☐	☐
I am in favour of using genetic or genomic testing in all disease areas to identify patients who have a disease, even before they have developed symptoms	☐	☐	☐
I am in favour of using genetic or genomic testing in all disease areas to identify patients who have a disease, and who are exhibiting symptoms for that disease	☐	☐	☐
I am in favour of using genetic or genomic testing in all disease areas to identify patients who will have a bad reaction and suffer complications when receiving the standard treatment for their disease	☐	☐	☐
I am in favour of using genetic or genomic testing to identify babies who have a genetic disorder, before they are born	☐	☐	☐
I am in favour of using genetic or genomic testing to identify babies who have a genetic disorder, just after the baby has been born	☐	☐	☐

QUESTION 21: YOUR CURRENT HEALTH STATUS

a) By placing a cross (X) in one box in each group below, please indicate which statements best describe your own health state today.

Mobility

I have no problems in walking about	☐
I have some problems in walking about	☐
I am confined to bed	☐

Self-Care

I have no problems with self-care	☐
I have some problems washing or dressing myself	☐
I am unable to wash or dress myself	☐

Usual Activities (*e.g. work, study, housework, family or leisure activities*)

I have no problems with performing my usual activities	☐
I have some problems with performing my usual activities	☐
I am unable to perform my usual activities	☐

Pain/Discomfort

I have no pain or discomfort	☐
I have moderate pain or discomfort	☐
I have extreme pain or discomfort	☐

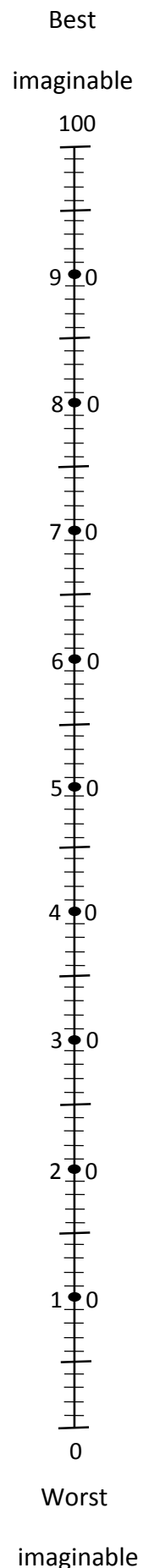
Anxiety/Depression

I am not anxious or depressed	☐
I am moderately anxious or depressed	☐
I am extremely anxious or depressed	☐

b) To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health is today, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is today.

**Your own
health state**



QUESTION 22

a) How did you find the questions in this survey?

☐

Easy

☐

Moderate

☐

Difficult

b) Are there any comments you would like to make regarding this questionnaire?

END OF QUESTIONNAIRE

Thank you for taking the time to complete this questionnaire.

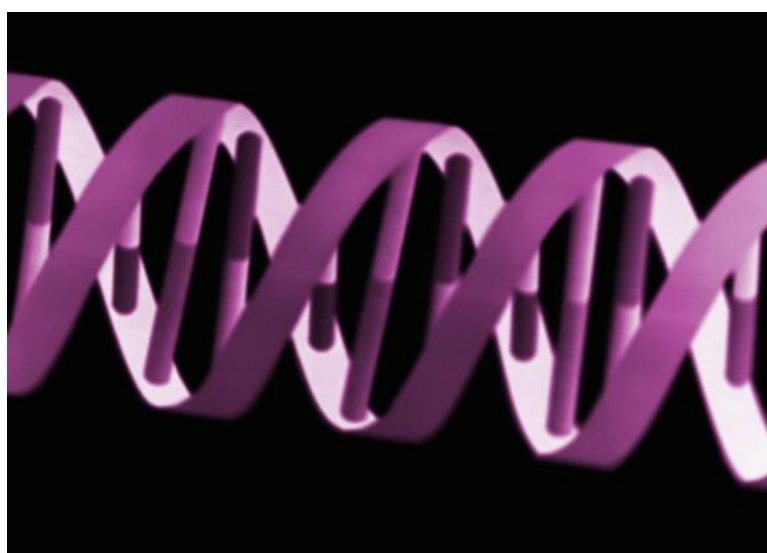
Please return your completed questionnaire to James Buchanan, or, if he is not around, please hand it in to the reception staff.

If you have not had time to complete the questionnaire today, please complete the questionnaire at home and send it to us by post free of charge, using the stamped addressed envelope that we have provided.

Appendix 32: Discrete choice experiment survey without risk graphics

Genomic testing and chronic lymphocytic leukaemia (CLL)

Do patients value tests which can predict their response to chemoimmunotherapy?



Doctors and researchers in Oxford are carrying out a survey about a new test which could be offered to chronic lymphocytic leukaemia (CLL) patients. We are interested in your views about how this new test could be provided by the NHS. Your participation is entirely voluntary, but your views would help us to develop this testing service in the future.

Please try to answer all of the questions in this booklet. There are no right answers - we are just interested in your views.

This questionnaire takes around 20 minutes to complete. Please return your completed questionnaire to James Buchanan, or, if he is not around, please hand it in to the reception staff. If you do not have time to complete the questionnaire today, please complete it at home and send it to us by post free of charge, using the stamped addressed envelope that we have provided.

If you have any questions about this questionnaire, please contact:

James Buchanan

NIHR Doctoral Research Fellow

Health Economics Research Centre, Department of Public Health
University of Oxford, Old Road Campus

Headington

Oxford

OX3 7LF

Telephone: 01865 289262 Email: oxfordcllsurvey@dph.ox.ac.uk

The doctor who is leading this project is **Dr Anna Schuh** (MD, PhD, MRCP, FRCPath) Consultant Haematologist, Oxford Cancer and Haematology Centre, Churchill Hospital, Headington, Oxford, OX3 7LE (Telephone: 01865 235259, Email: anna.schuh@nhs.net).

For more information about this study, visit: www.herc.ox.ac.uk/ctl

HOW TO COMPLETE THIS QUESTIONNAIRE

We would like you to imagine that you are a CLL patient who requires chemoimmunotherapy treatment (e.g. combined fludarabine, cyclophosphamide, and rituximab (FCR) treatment). Out of every **400** CLL patients, we expect that about **300** will respond really well to the treatment that is usually tried first, and **100** will not respond (or will relapse within 2 years of treatment). **There is no way of knowing which group you are in before you start treatment.**

If we could identify the patients that will not respond to the usual chemoimmunotherapy treatment before this treatment is started, we might be able to avoid giving them treatment that will not work: an alternative treatment would be offered instead.

Next, we would also like you to imagine that you have been offered a test which could provide you and your doctors with some information about how you will respond to the usual treatment. This test is called a **genomic test.**

This genomic test works by identifying variations in the genetic make-up (DNA) of your leukaemia cells. These variations affect how you will respond to the usual chemoimmunotherapy treatment. This information helps your doctors to decide whether to give you the **usual chemoimmunotherapy treatment** or whether to give you an **alternative chemoimmunotherapy treatment.** This test **does not** look at inheritable genetic information, it only looks at the genes in the leukaemia cells.

In this survey you will be presented with several **imaginary** situations in which two different tests will be described. In each situation, we would like you to choose which test you prefer out of two options: **Test A** or **Test B.** Even if you do not like either test, please mark the one you would prefer if you had to choose one.

Each imaginary test will be described to you in terms of six features:

- Time to receive the test result
- Cost of the test
- Ability of the test to predict who will not respond to the usual chemotherapy treatment
- Test reliability
- Length of time health care professionals spend describing the test and the test results to you
- Type of health care professional who explains the test result to you

A detailed description of each feature is on the next page. Please read these descriptions carefully, then try to answer all of the questions that follow.

There are no right or wrong answers - we are just interested in your views. All information will be kept strictly confidential and used for research purposes only.

THANK YOU FOR YOUR TIME AND ASSISTANCE

FEATURES

Each imaginary test will be described to you in terms of the following features:

- **Time to receive test result**

This describes how many days are between you providing a blood sample and receiving the test result. You will only begin a chemoimmunotherapy treatment once you receive the test result.

- **Cost of test to you**

This describes how much it costs for one test. Obviously in the NHS you are not asked to pay for medical tests – the NHS covers this cost. However, for the purposes of this survey, imagine that you live in a country where you have to cover the cost of your medical tests out of your own pocket, such as the United States of America (USA).

- **Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment**

Out of 400 CLL patients like you, we expect that about 300 will respond really well to the chemoimmunotherapy treatment that is usually tried first, and 100 will not respond (or will relapse within 2 years of treatment). Some genomic tests can detect most of the 100 patients who will not respond. Other genomic tests can only detect a few of these patients. The ability of the genomic test to predict if you are one of the patients who will not respond to usual treatment is given by this feature.

A figure of 30 means that the genomic test will pick up 30 out of every 100 CLL patients who will not respond to usual treatment.

A figure of 90 means that the genomic test will pick up 90 out of every 100 patients who will not respond to usual treatment.

- **Test reliability**

Occasionally a genomic test will provide an incorrect result: the test result suggests that you won't respond to the usual chemoimmunotherapy treatment when in fact you will. This means that you will incorrectly receive an alternative treatment instead of the usual treatment. This alternative treatment may be slightly less effective than the usual treatment. The most reliable tests rarely provide an incorrect result. Less reliable tests return incorrect results more often. The reliability of the genomic test is given by this feature.

A figure of 2 means that for every 100 patients who have the test, 2 will get an incorrect result and will be given an alternative treatment instead of the usual treatment.

A figure of 8 means that for every 100 patients who have the test, 8 will get an incorrect result and will be given an alternative treatment instead of the usual treatment.

- **Length of time health care professionals spend describing the test and the test results to you**

5 minutes: Before the test, your health care professional tells you that they will be taking a blood sample in order to do a genomic test that will provide information on whether you will respond to the usual chemoimmunotherapy treatment. After you have had the test, he/she tells you whether the test has indicated that you will respond to the usual treatment, and recommends an alternative treatment if necessary.

10 minutes: In addition to the information provided in 5 minutes, your health care professional also provides some limited information about how the test works and the different chemoimmunotherapy treatments.

15 minutes: In addition to the information provided in 10 minutes, your health care professional also describes in detail the consequences of different chemoimmunotherapy treatments in terms of side-effects, success rates and prognosis.

20 minutes: In addition to the information provided in 15 minutes, your health care professional also provides detailed test results (explaining which genetic mutations have been found, if any) and has time to answer any questions you may have about any potential consequences for your family.

- **Type of health care professional who explains the test result to you**

General practitioner: This is the GP who you would normally go to see in your local doctors surgery or health centre.

Specialist nurse: This is a nurse who specialises in CLL and who works in the outpatient clinic of the hospital that you normally go to for check-ups and treatments.

Junior hospital doctor: This is a doctor who works in the outpatient clinic of the hospital that you normally go to for check-ups and treatments. He or she has some knowledge of CLL.

Consultant hospital doctor: This is a doctor who works in the outpatient clinic of the hospital that you normally go to for check-ups and treatments. He or she is an experienced CLL specialist.

DESCRIBING A GENOMIC TEST

Each imaginary test will always be described to you in terms of the same six features. Each feature has four possible levels. These levels will change from one question to the next. Please read through the descriptions of the features and levels.

Feature	Possible levels
Time to receive test result	5 days
	8 days
	11 days
	14 days

Cost of test to you (remember, in the NHS you would not have to actually pay for these tests)	£130
	£260
	£400
	£600

Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment
	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment
	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment
	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment

Test reliability (number of incorrect results provided every 100 tests)	<u>2</u> out of every <u>100</u> tests provide an incorrect result
	<u>4</u> out of every <u>100</u> tests provide an incorrect result
	<u>6</u> out of every <u>100</u> tests provide an incorrect result
	<u>8</u> out of every <u>100</u> tests provide an incorrect result

Length of time health care professionals spend describing the test and the test results to you	5 minutes
	10 minutes
	15 minutes
	20 minutes

Type of health care professional who explains the test result to you	General practitioner
	Specialist nurse
	Junior hospital doctor
	Consultant hospital doctor

Once you have read through the descriptions, we would like to know which of these features are most important to you. Please complete **QUESTION 1** on the next page.

QUESTION 1: FEATURE RANKING EXERCISE

Please consider all 6 of the features that make up the test and rank these in order of importance to you in the yellow boxes.

Rank the feature that is most important to you with a 1. Then rank the second most important feature with a 2. Do the same for the third, fourth and fifth most important features. Rank the feature that is least important to you with a 6.

Feature	Rank
Time to receive the test result	
Cost of the test <u>to you</u>	
Ability of the test to predict who will <u>not</u> respond to the usual chemoimmunotherapy treatment	
Test reliability (number of incorrect results provided every 100 tests)	
Length of time health care professionals spend describing the test and the test results to you	
Type of health care professional who explains the test result to you	

Once you have ranked these features, read through the **EXAMPLE QUESTION** on the next page. This will provide some guidance on answering the questions.

Then go to **QUESTION 2**.

Indicate with a cross (X) whether you would prefer **TEST A** or **TEST B**. You can do this by writing X in the relevant yellow box.

Repeat this for **QUESTIONS 3 to 18**.

EXAMPLE QUESTION – DO NOT ANSWER

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	5 days
Cost of the test to you	£260	£600
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result	<u>4</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	10 minutes	5 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	Junior hospital doctor

Which test would you prefer?

☐

TEST A

☒

TEST B

QUESTION 2

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	5 days
Cost of the test to you	£260	£130
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>6</u> out of every <u>100</u> tests provide an incorrect result	<u>2</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	10 minutes	15 minutes
Type of health care professional who explains the test result to you	General practitioner	Consultant hospital doctor

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 3

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	5 days	14 days
Cost of the test to you	£130	£600
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>8</u> out of every <u>100</u> tests provide an incorrect result	<u>8</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	5 minutes	10 minutes
Type of health care professional who explains the test result to you	Specialist nurse	Junior hospital doctor

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 4

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	5 days	14 days
Cost of the test to you	£400	£130
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>8</u> out of every <u>100</u> tests provide an incorrect result	<u>2</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	10 minutes	15 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	General practitioner

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 5

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	14 days	5 days
Cost of the test to you	£260	£400
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>6</u> out of every <u>100</u> tests provide an incorrect result	<u>2</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	10 minutes	20 minutes
Type of health care professional who explains the test result to you	Specialist nurse	Junior hospital doctor

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 6

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	11 days
Cost of the test to you	£600	£600
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result	<u>8</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	10 minutes	20 minutes
Type of health care professional who explains the test result to you	General practitioner	Specialist nurse

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 7

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	5 days	14 days
Cost of the test to you	£400	£260
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>4</u> out of every <u>100</u> tests provide an incorrect result	<u>6</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	20 minutes	5 minutes
Type of health care professional who explains the test result to you	Specialist nurse	General practitioner

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 8

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	14 days	5 days
Cost of the test to you	£260	£400
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>6</u> out of every <u>100</u> tests provide an incorrect result	<u>6</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	20 minutes	15 minutes
Type of health care professional who explains the test result to you	Junior hospital doctor	Consultant hospital doctor

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 9

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	14 days	5 days
Cost of the test to you	£130	£400
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>8</u> out of every <u>100</u> tests provide an incorrect result	<u>4</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	20 minutes	5 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	Specialist nurse

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 10

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	14 days	5 days
Cost of the test to you	£600	£260
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>50</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result	<u>4</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	5 minutes	20 minutes
Type of health care professional who explains the test result to you	Junior hospital doctor	General practitioner

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 11

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	5 days	14 days
Cost of the test to you	£400	£600
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>6</u> out of every <u>100</u> tests provide an incorrect result	<u>4</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	15 minutes	20 minutes
Type of health care professional who explains the test result to you	Junior hospital doctor	Specialist nurse

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 12

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	11 days	8 days
Cost of the test to you	£600	£130
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>4</u> out of every <u>100</u> tests provide an incorrect result	<u>2</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	15 minutes	10 minutes
Type of health care professional who explains the test result to you	General practitioner	Junior hospital doctor

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 13

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	11 days	8 days
Cost of the test to you	£260	£400
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>4</u> out of every <u>100</u> tests provide an incorrect result	<u>6</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	10 minutes	5 minutes
Type of health care professional who explains the test result to you	Specialist nurse	Consultant hospital doctor

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 14

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	11 days
Cost of the test to you	£130	£130
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result	<u>8</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	5 minutes	15 minutes
Type of health care professional who explains the test result to you	General practitioner	Specialist nurse

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 15

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	11 days
Cost of the test to you	£600	£600
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>8</u> out of every <u>100</u> tests provide an incorrect result	<u>2</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	15 minutes	10 minutes
Type of health care professional who explains the test result to you	Junior hospital doctor	Consultant hospital doctor

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 16

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	8 days	11 days
Cost of the test to you	£400	£260
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>70</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>6</u> out of every <u>100</u> tests provide an incorrect result	<u>4</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	20 minutes	15 minutes
Type of health care professional who explains the test result to you	General practitioner	Consultant hospital doctor

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 17

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	11 days	8 days
Cost of the test to you	£260	£260
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>4</u> out of every <u>100</u> tests provide an incorrect result	<u>6</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	15 minutes	5 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	General practitioner

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 18

Consider the following features describing two tests (**Test A** or **Test B**). Please indicate with a cross (X) which test you would choose:

Feature	Test A	Test B
Time to receive the test result	11 days	8 days
Cost of the test to you	£130	£130
Ability of the test to predict who will not respond to the usual chemoimmunotherapy treatment	Test identifies <u>30</u> out of every <u>100</u> patients who will not respond to usual treatment	Test identifies <u>90</u> out of every <u>100</u> patients who will not respond to usual treatment
Test reliability	<u>2</u> out of every <u>100</u> tests provide an incorrect result	<u>8</u> out of every <u>100</u> tests provide an incorrect result
Length of time health care professionals spend describing the test and the test results to you	5 minutes	10 minutes
Type of health care professional who explains the test result to you	Consultant hospital doctor	Junior hospital doctor

Which test would you prefer?

☐

TEST A

☐

TEST B

QUESTION 19: BACKGROUND INFORMATION

We would now like to ask some questions about **you**. All the information that you provide will help us in our analysis, and all of your details will be kept confidential. If you do not wish to answer some of these questions you do not have to.

a) What is your gender?

☐

Male

☐

Female

b) What is your age?

years old

c) What is your occupational status?

☐

Employed full time

☐

Employed part-time

☐

Self-employed

☐

Unemployed

☐

Retired

☐

Homemaker/Stay at home parent

☐

Student

☐

Other

d) How old were you when you left full time education?

years old

e) Please estimate the annual income of your household before tax and national insurance:

☐

Less than £10,000

☐

£10,000 - £19,999

☐

£20,000 - £29,999

☐

£30,000 - £39,999

☐

£40,000 and over

f) How many people live in your household?

people

g) How many children do you have?

Living at home

Living away from home

h) How long does it take you to reach the hospital that you usually attend for treatment?

☐

0-15 minutes

☐

16-30 minutes

☐

31-45 minutes

☐

46-60 minutes

☐

61-75 minutes

☐

76-90 minutes

☐

More than 90 minutes

i) What mode of transport do you normally take to reach the hospital that you usually attend for treatment? Please select the mode on which you complete most of your journey.

☐

Hospital arranged transport (e.g. ambulance, hospital arranged taxi)

☐

Public transport (e.g. privately arranged taxi, bus, train)

☐

Private transport (e.g. private car, motorcycle, bicycle, walk)

j) How long ago were you diagnosed with CLL, to the nearest year?

years

k) Have you ever undergone chemoimmunotherapy treatment for your CLL?

☐

Yes (go to question 1)

☐

No (go to question 20)

l) If yes, how many times have you undergone a course of chemoimmunotherapy treatment? If you have received one type of chemoimmunotherapy several times over a period of months, this counts as one course of treatment.

times

m) Have you ever stayed in hospital overnight as a result of undergoing chemoimmunotherapy treatment? ☐ Yes

☐ No

n) Are you currently undergoing chemoimmunotherapy treatment for your CLL? ☐ Yes (go to question o)

☐ No (go to question 20)

o) If yes, what type of treatment are you currently receiving?

QUESTION 20: GENETIC AND GENOMIC TESTING

a) Do you know if you have had a genetic or genomic test since being diagnosed with CLL (sometimes these are called FISH tests, or tests to see if you have a certain gene mutation)? ☐ Yes (go to question b)

☐ No (go to question d)

☐ Not sure (go to question d)

b) If yes, do you know what type of genetic test this was?

c) Did you find the information provided by this test to be useful?

☐ Yes

☐ No

☐ Not sure

d) Do you know anybody (e.g. a family member or a friend) who has had a genetic or genomic test? ☐ Yes

☐ No

e) Please indicate whether you agree or disagree with each of these statements. Place a cross (X) by either Yes, No or Not sure for each statement.

Statement	Yes	No	Not sure
I am in favour of using genetic or genomic testing to identify CLL patients who will not respond to usual chemoimmunotherapy treatment	☐	☐	☐
I am in favour of using genetic or genomic testing to identify patients with other diseases who will not respond to usual treatments	☐	☐	☐
I am in favour of using genetic or genomic testing in all disease areas to identify patients who have a disease, even before they have developed symptoms	☐	☐	☐
I am in favour of using genetic or genomic testing in all disease areas to identify patients who have a disease, and who are exhibiting symptoms for that disease	☐	☐	☐
I am in favour of using genetic or genomic testing in all disease areas to identify patients who will have a bad reaction and suffer complications when receiving the standard treatment for their disease	☐	☐	☐
I am in favour of using genetic or genomic testing to identify babies who have a genetic disorder, before they are born	☐	☐	☐
I am in favour of using genetic or genomic testing to identify babies who have a genetic disorder, just after the baby has been born	☐	☐	☐

QUESTION 21: YOUR CURRENT HEALTH STATUS

a) By placing a cross (X) in one box in each group below, please indicate which statements best describe your own health state today.

Mobility

I have no problems in walking about	☐
I have some problems in walking about	☐
I am confined to bed	☐

Self-Care

I have no problems with self-care	☐
I have some problems washing or dressing myself	☐
I am unable to wash or dress myself	☐

Usual Activities (*e.g. work, study, housework, family or leisure activities*)

I have no problems with performing my usual activities	☐
I have some problems with performing my usual activities	☐
I am unable to perform my usual activities	☐

Pain/Discomfort

I have no pain or discomfort	☐
I have moderate pain or discomfort	☐
I have extreme pain or discomfort	☐

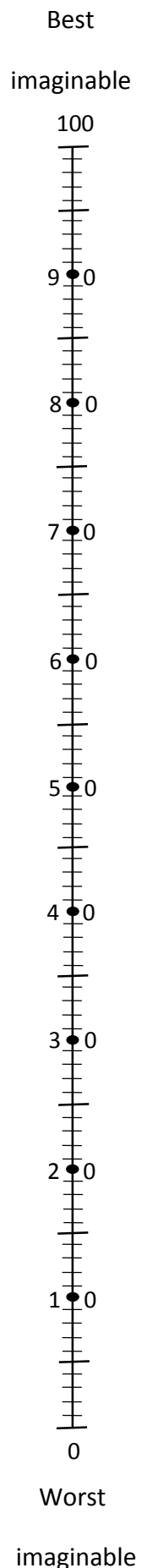
Anxiety/Depression

I am not anxious or depressed	☐
I am moderately anxious or depressed	☐
I am extremely anxious or depressed	☐

b) To help people say how good or bad a health state is, we have drawn a scale (rather like a thermometer) on which the best state you can imagine is marked 100 and the worst state you can imagine is marked 0.

We would like you to indicate on this scale how good or bad your own health is today, in your opinion. Please do this by drawing a line from the box below to whichever point on the scale indicates how good or bad your health state is today.

**Your own
health state**



QUESTION 22

a) How did you find the questions in this survey?

☐

Easy

☐

Moderate

☐

Difficult

b) Are there any comments you would like to make regarding this questionnaire?

END OF QUESTIONNAIRE

Thank you for taking the time to complete this questionnaire.

Please return your completed questionnaire to James Buchanan, or, if he is not around, please hand it in to the reception staff.

If you have not had time to complete the questionnaire today, please complete the questionnaire at home and send it to us by post free of charge, using the stamped addressed envelope that we have provided.

Appendix 33: Discrete choice experiment respondent opinions on genetic testing and attribute importance rankings

In section four of the discrete choice experiment survey, respondents were asked for their opinions on genetic or genomic testing. Seven clinical contexts were described and respondents were asked whether they were in favour of genetic testing being used in each context. Possible answers were 'Yes', 'No' and 'Not sure', with only an answer of 'Yes' taken as agreement with a statement. Respondents were generally positive about testing. Over 90% of respondents agreed with four of the seven statements, and over 50% agreed with six of the seven statements.

Table 1: Respondent opinions concerning genetic and genomic testing

Genetic testing statement	% who agree (n=217)
1. I am in favour of using genetic or genomic testing to identify CLL patients who will not respond to usual chemoimmunotherapy treatment	96.3%
2. I am in favour of using genetic or genomic testing to identify patients with other diseases who will not respond to usual treatments	94.0%
3. I am in favour of using genetic or genomic testing in all disease areas to identify patients who have a disease, even before they have developed symptoms	59.9%
4. I am in favour of using genetic or genomic testing in all disease areas to identify patients who have a disease, and who are exhibiting symptoms for that disease	91.2%
5. I am in favour of using genetic or genomic testing in all disease areas to identify patients who will have a bad reaction and suffer complications when receiving the standard treatment for their disease	95.4%
6. I am in favour of using genetic or genomic testing to identify babies who have a genetic disorder, before they are born	45.2%
7. I am in favour of using genetic or genomic testing to identify babies who have a genetic disorder, just after the baby has been born	58.8%

CLL = chronic lymphocytic leukaemia.

Table 2: Average attribute importance rankings, prior to undertaking the discrete choice experiment

Attribute	Mean (SD) ^a
Ability of the test to predict who will respond to the usual chemotherapy treatment [EFFECT]	1.9 (1.2)
Test reliability [REL]	2.2 (1.2)
Time to receive the test result [TIME]	3.9 (1.6)
Length of time clinicians spend describing the test to you [INFO]	3.9 (1.2)
Type of clinician who explains the test result to you [WHO]	4.1 (1.4)
Cost of the test [COST]	4.9 (1.3)

^a 1 = most important; 6 = least important.

Appendix 34: Incorporating attribute interactions into Model A

Table 1: Incorporating attribute interactions into Model A (multinomial logit regression)

Attribute	Model A (main effects, MNL)		Model A (including respondent characteristics, MNL)	
	Coefficients	WTP (95% CI)	Coefficients	WTP (95% CI)
<i>Main effects^a</i>				
TIME	-0.0461**	-£27.19 (-£38.67 to -£15.71)	-0.1077**	-£70.66 (-£118.61 to -£22.72)
COST	-0.0017**	N/A	-0.0015**	N/A
EFFECT	0.0404**	£23.79 (£16.06 to £31.51)	0.0259**	£17.03 (£6.90 to £27.15)
REL	-0.1584**	-£93.38 (-£125.79 to -£60.98)	-0.2103**	-£137.98 (-£193.11 to -£82.84)
INFO0	-0.1367*	-£80.59 (-£165.73 to £4.54)	-0.2714**	-£178.07 (-£315.02 to -£41.12)
INFO1	-0.0319	-£18.83 (-£81.66 to £43.99)	-0.0305	-£20.03 (-£93.91 to £53.84)
INFO2	0.2624**	£154.69 (£55.61 to £253.76)	0.2848**	£186.88 (£63.74 to £310.01)
INFO3 ^b	-0.0937**	-£55.26	0.0171**	£11.22
WHO0	-0.8399**	-£495.10 (-£646.56 to -£343.64)	-0.9531**	-£625.40 (-£856.08 to -£394.73)
WHO1	0.5071**	£298.92 (£181.14 to £416.71)	0.5416**	£355.40 (£198.52 to £512.28)
WHO2	-0.2588**	-£152.54 (-£251.26 to -£53.81)	-0.3550**	-£232.96 (-£375.84 to -£90.09)
WHO3 ^b	0.5915**	£348.71	0.7665**	£502.97
<i>Interaction terms</i>				
MALE x INFO0	-	-	0.2174*	£142.68 (£20.27 to £265.09)
MALE x WHO0	-	-	0.2177*	£142.87 (£5.02 to £280.72)
INCOME x REL	-	-	0.0798**	£52.39 (£14.80 to £89.99)
DIAGNOSIS ^c x EFFECT	-	-	0.0012**	£0.80 (£0.25 to £1.35)
CHEMO ^d x EFFECT	-	-	-0.0140**	-£9.19 (-£14.09 to -£4.28)
GENETIC TEST ^e x WHO2	-	-	0.2758**	£180.94 (£33.67 to £328.22)
EQ-5D x TIME	-	-	0.0007*	£0.48 (-£0.03 to £0.98)
EQ-5D x EFFECT	-	-	0.0002**	£0.12 (£0.02 to £0.21)
<i>Number of choice tasks</i>	3 097	-	2 818	-
<i>Pseudo-R²</i>	0.4101	-	0.4253	-

95% CI = 95% confidence interval; MNL = multinomial logit; WTP = willingness-to-pay. * P-value < 0.05, ** P-value < 0.01. ^a The coefficient for the alternative specific constant (ASC) was not significant in the initial regressions (indicating no respondent left-right bias) so was dropped from subsequent regressions. ^b WTP 95% CIs cannot be calculated for the base levels of the effects coded variables. ^c Time since diagnosis. ^d Ever had chemotherapy. ^e Previously undergone genetic testing.

Table 2: Alternative specifications for the version of Model A that includes attribute interactions

Specification	AIC	BIC	CAIC
Multinomial logit	2,281	2,400	N/A
Alternative-specific conditional logit	2,283	2,409	N/A
Mixed logit	1,930	2,169	N/A
Latent class (three classes)	1,842	1,842	2,187
Generalised multinomial logit	Did not converge		N/A

AIC = Akaike information criterion; BIC = Bayesian information criterion; CAIC = Consistent information criterion.

The results suggest that a three-class latent class model is the most appropriate specification for the version of Model A that includes attribute interactions. The regression results for this model are presented in the table below.

Table 3: Regression results for the three-class latent class specification of Model A (including attribute interactions)

Attribute	Coefficients		
	Class One	Class Two	Class Three
TIME	-0.185	0.089	0.390
COST	-0.001	-0.002	-0.013
EFFECT	0.010	0.163	-2.021
REL	-0.300	-0.438	-1.205
INFO0	-0.206	-0.968	-1.547
INFO1	-0.106	0.156	-2.743
INFO2	0.059	0.234	3.460
WHO0	-0.659	-2.002	-17.673
WHO1	0.258	1.315	9.509
WHO2	-0.498	-0.158	-10.033
MALE x INFO0	0.315	0.328	-3.499
MALE x WHO0	0.034	0.417	4.569
INCOME x REL	0.226	-0.051	0.255
DIAGNOSIS ^a x EFFECT	0.002	0.002	0.002
CHEMO ^b x EFFECT	-0.021	-0.015	0.164
GENETIC TEST ^c x WHO2	0.656	-0.032	4.462
EQ-5D x TIME	0.002	-0.002	-0.014
EQ-5D x EFFECT	0.000	-0.001	0.030
Class share	30%	57%	14%

^a Time since diagnosis; ^b Ever had chemotherapy; ^c Previously undergone genetic testing

The multinomial logit (MNL) version of Model A (including interactions) indicates that men place a higher value on shorter appointments and receiving test results from a general practitioner. This suggests that respondents in class two are more likely to be male. The same MNL version of Model A also indicates that those with higher incomes place a lower valuation on a decrease in test reliability, which suggests that respondents in classes one and three are likely to have higher incomes. By the same reasoning, classes one and two are more likely to contain respondents who have previously undergone chemotherapy, classes one and three are more likely to contain respondents who have previously had a genetic test, and class two is more likely to contain respondents in poor health. Putting this information together with the other coefficient values suggests the following respondent groups:

- **Class One:** High income patients, who have previously had both chemotherapy and a genetic test. These patients are unwilling to wait for test results and prefer more effective tests, but

are less concerned about how much information they receive and who delivers their test results. They could be characterised as demanding patients whose previous treatment experiences have generated strong opinions about future treatment options.

- Class Two: Low income male patients, in poor health, who have previously had chemotherapy but no genetic test. These patients care less about timeliness of results and test reliability, but have strong preferences for effective tests. Given their health status, they may have failed chemotherapy previously, so value a test that can predict future treatment response.
- Class Three: High income patients who haven't previously had chemotherapy, but have had a genetic test. These patients have several counter-intuitive preferences (e.g. prefer slower test results and less effective tests). Given this, and the small size of this group, no further conclusions are drawn about the characteristics of these patients.

Appendix 35: Sensitivity analysis for the cost-benefit analysis

Table 1: Univariate sensitivity analysis results for the cost-benefit analysis

Parameter	Variation	Cost		WTP		Net benefit		Net benefit difference	Test with highest net benefit
		Genetic test	Genomic test	Genetic test	Genomic test	Genetic test	Genomic test		
Base case	-	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
TIME	Low	£281.08	£319.23	-£3,184.81	-£2,023.06	-£3,465.89	-£2,342.29	£1,123.60	Genomic
	High	£281.08	£319.23	-£2,734.64	-£1,572.89	-£3,015.72	-£1,892.12	£1,123.60	Genomic
EFFECT	Low	£281.08	£319.23	-£3,106.87	-£2,160.38	-£3,387.95	-£2,479.61	£908.34	Genomic
	High	£281.08	£319.23	-£2,812.58	-£1,435.56	-£3,093.66	-£1,754.79	£1,338.87	Genomic
REL	Low	£281.08	£319.23	-£4,389.83	-£2,968.06	-£4,670.91	-£3,287.29	£1,383.62	Genomic
	High	£281.08	£319.23	-£1,529.62	-£627.89	-£1,810.70	-£947.12	£863.59	Genomic
INFO0	Low	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
	High	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
INFO1	Low	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
	High	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
INFO2	Low	£281.08	£319.23	-£3,074.11	-£1,912.35	-£3,355.19	-£2,231.58	£1,123.60	Genomic
	High	£281.08	£319.23	-£2,845.35	-£1,683.59	-£3,126.43	-£2,002.82	£1,123.60	Genomic
INFO3	Low	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
	High	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
WHO0	Low	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
	High	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
WHO1	Low	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
	High	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
WHO2	Low	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
	High	£281.08	£319.23	-£2,959.73	-£1,797.97	-£3,240.81	-£2,117.20	£1,123.60	Genomic
WHO3	Low	£281.08	£319.23	-£3,109.73	-£1,947.97	-£3,390.81	-£2,267.20	£1,123.60	Genomic
	High	£281.08	£319.23	-£2,809.73	-£1,647.97	-£3,090.81	-£1,967.20	£1,123.60	Genomic
Genetic test cost ^a	Low	£140.54	£319.23	-£2,959.73	-£1,797.97	-£3,100.27	-£2,117.20	£983.06	Genomic
	High	£421.62	£319.23	-£2,959.73	-£1,797.97	-£3,381.35	-£2,117.20	£1,264.14	Genomic
Genomic test cost ^a	Low	£281.08	£159.62	-£2,959.73	-£1,797.97	-£3,240.81	-£1,957.59	£1,283.22	Genomic
	High	£281.08	£478.85	-£2,959.73	-£1,797.97	-£3,240.81	-£2,276.82	£963.99	Genomic
Improved genomic testing ^b	-	£281.08	£319.23	-£2,959.73	£1,326.09	-£3,240.81	£1,006.86	£4,247.67	Genomic

Parameter	Variation	Cost		WTP		Net benefit		Net benefit difference	Test with highest net benefit
		Genetic test	Genomic test	Genetic test	Genomic test	Genetic test	Genomic test		
Patient sees junior hospital doctor	-	£281.08	£319.23	-£3,656.40	-£2,494.65	-£3,937.48	-£2,813.88	£1,123.60	Genomic
Patient sees general practitioner	-	£281.08	£319.23	-£3,991.00	-£2,829.24	-£4,272.08	-£3,148.47	£1,123.60	Genomic
Genomic test appointment extended to 20 minutes	-	£281.08	£319.23	-£2,959.73	-£1,876.26	-£3,240.81	-£2,195.49	£1,045.32	Genomic
Genomic test results returned in 10 days	-	£281.08	£319.23	-£2,959.73	-£1,673.77	-£3,240.81	-£1,993.00	£1,247.81	Genomic
Risk graphics used	-	£281.08	£319.23	-£1,712.41	-£693.06	-£1,993.49	-£1,012.29	£981.20	Genomic
Risk graphics not used	-	£281.08	£319.23	-£3,118.60	-£2,045.03	-£3,399.68	-£2,364.26	£1,035.42	Genomic

WTP = willingness-to-pay. ^a Low and high values based on values used in one-way sensitivity analysis in the cost-effectiveness and cost-utility analysis. ^b Test effectiveness changed to 25% (i.e. 25% of patients identified as FCR non-responders) and reliability changed to 1%.

Appendix 36: Extending the extra-welfarist economic evaluation to a societal perspective – methods and results

This appendix describes the various steps involved in extending the extra-welfarist economic evaluation from a NHS perspective to a societal perspective. The results of this extension to the analysis are also described in full.

Methods

The costs that were included in this analysis fell into three categories. The steps involved in estimating these costs are described below. All of the parameters that informed the societal cost calculations are reported in **Table 1**.

Productivity costs

These costs are incurred when the productivity of chronic lymphocytic leukaemia (CLL) patients is affected by illness, treatment, disability or premature death [92]. Although many CLL patients are elderly, a proportion of this population are still employed, so these costs are still potentially important. In addition to the costs that were captured in the five sub-categories below, it is also possible that compensation mechanisms and multiplier effects exist which act on these estimates. However, methods to account for these effects are currently underdeveloped [246,359] so they are not included in the analysis.

Short-term absenteeism due to CLL-related morbidity

Short-term absences can occur due to time taken off work (or leisure) to attend routine healthcare appointments, or as a consequence of treatment. Such absences either result in lost earnings (due to time taken off work) or lost leisure time, which also has a monetary value. UK government data from 2013 indicates that the proportion of the population in employment is 80.9% (25-49 year olds), 67.4% (50-64 year olds) and 9.5% (65 years and older) [360]. As the starting age of patient cohorts in the Markov model is 65 years in the base case analysis, the estimate of 9.5% was applied. However, this proportion will diminish over time as patients age in the model and retire from employment. In the absence of data on the rate of change, this proportion was reduced by 20% a year such that no patients in the model were working five years after retirement (i.e. after reaching 70 years in age). This assumption was varied between one year and ten years in sensitivity analysis.

The costs associated with short-term absenteeism were then estimated for both employed and unemployed (or, more likely, retired) patients. For employed patients, data were available from a German study on events experienced by CLL patients which can lead to short-term absenteeism. This study reported that on average each year CLL patients take 7.6 sick days, have 14.7 physician contacts, and spend 8.9 days in hospital [295]. The simplifying assumption was made that sick days and days in hospital could only be experienced during symptomatic states in the Markov model i.e. not in the following states:

- General remission states;
- Bone marrow transplant (BMT – it was assumed that patients undergoing BMT would require the whole month off work as treatment is intensive);
- BMT remission states;

- Receiving best supportive care (BSC – it was assumed that patients would be off work permanently in this state);
- Receiving ibrutinib treatment (as patients receive ibrutinib until they become symptomatic).

The number of sick days or inpatient days per 28-day cycle in the model was therefore estimated to be 1.27. Physician contacts were assumed to require half a day away from employment, so the number of days of physician contacts per 28-day cycle was estimated to be 0.56. As these are only short-term absences from employment, either the human capital or friction cost approach can be used to value these absences [92]. Costs were therefore calculated using data on average UK hourly wage rates, adjusted for age, and assuming that a whole day at work is 7.5 hours [361].

Patients who are unemployed (or retired) were assumed to experience the same events as those in employment, in the same model states. However, the valuation of these events differed: minimum wage rates were applied to reflect the opportunity cost of the time of these patients (as per recently published cost of illness studies in heart disease and cancer [362-365]).

Short-term presenteeism due to CLL-related morbidity

Presenteeism captures the impact of reduced productivity during work or leisure time. There is little data available on the costs associated with presenteeism, either in CLL or cancer more generally. One literature review from 2008 found only one study reporting presenteeism rates for cancer [366]. This study reported that presenteeism in cancer is 8.45% over the course of a year, hence 2.4 days are lost per 28-day model cycle [367]. This percentage was applied during symptomatic model states as per the short-term absenteeism calculations, and average UK hourly wage rates (or minimum wage rates) were applied to value this time [361,365]. This approach assumes that employees experiencing presenteeism are 0% effective during this time, which is a standard assumption in analyses of presenteeism. This assumption was tested in sensitivity analysis.

Long-term absenteeism due to CLL-related morbidity

No data was identified on long-term absenteeism in CLL. It was therefore assumed that this only occurs in the BSC state, until retirement age (assumed to be 65 years for both men and women). Information was extracted from the Markov model on the average number of cycles spent in BSC for each comparator, and average UK hourly wage rates were applied to value this time [361]. Current guidance recommends that the friction cost approach is applied to value long-term absenteeism [294], so only absences up to 90 days were included in calculations. As the starting age of patients in the model is 65 years, there are no costs associated with long-term absenteeism in the base case analysis. However, starting ages of 25 years and 45 years were evaluated in a sensitivity analysis and in these scenarios a proportion of the patient cohort experienced long-term absenteeism.

Productivity losses due to CLL-related mortality

Information was extracted from the model on the number of deaths due to CLL in each cycle. This data was combined with information on retirement age to calculate working years lost to CLL, which were then valued using average UK hourly wage rates [361]. Again, there are no costs associated with productivity losses in the base case analysis due to the assumed starting age of 65 years.

Productivity gains due to improved treatment

Productivity gains due to improved treatment reflect additional healthy time used for both work and leisure. It has been noted previously that the value of additional healthy time used for leisure is probably captured in the utility scores for health states [12]. Capturing the value of additional healthy time used for work is more problematic, and primarily depends on whether patients considered this when providing utility data. No information is available in the studies from which utility data was extracted for the cost-utility analysis (CUA) on whether patients were asked to consider the implications of their choices for their ability to work [239]. It is therefore unclear whether the value of additional healthy time used for work has already been captured. A number of papers suggest that productivity effects mainly impact on costs rather than outcomes, but evidence is generally lacking on this effect [368-370]. For the purposes of this analysis it was assumed that any productivity gains due to improved treatment were captured by the utility scores. Any missed productivity gains are likely to be small in size as many patients are either retired or approaching retirement.

Informal care costs

CLL patients frequently need a large amount of informal care in addition to the care provided by healthcare professionals. Information on the time costs of informal caregiving were extracted from a study which estimated the economic burden of blood disorders and blood cancers across the European Union (EU) [371]. This study provided information on the number of hours of informal care provided to two patient groups. The first group was patients dying with blood cancer. For the purposes of this analysis, these patients were assumed to be those receiving BSC. The second patient group was patients severely hampered by blood cancer, who were assumed to be those in any other symptomatic health state. In both cases, information was available on the number of hours of care provided by both employed and non-employed informal caregivers. Information on average UK hourly wage rates (or minimum wage rates) was applied to value this time and costs per 28-day model cycle were calculated [361,365].

Out of pocket costs

Out of pocket (OOP) costs can be incurred by both CLL patients and informal caregivers.

Patients

Information on the OOP costs of patients was extracted from a report on the costs associated with cancer published by Macmillan, a UK cancer charity [296]. This report contained the results of a survey conducted amongst 1,600 UK cancer patients in 2012 which captured self-reported expenditure across a range of areas over a six month period. Categories of expenditure included costs associated with receiving outpatient and inpatient care (e.g. travel costs), costs associated with day-to-day living (e.g. home help), other healthcare costs (e.g. over the counter medicines), and clothing and equipment costs. Information was extracted from the report on all categories of expenditure except for the costs of help around the home and personal care, to avoid double counting with informal care time costs. Average costs per 28-day cycle were then calculated. It should be noted that this publication did not report the costs associated with admission to a nursing home. These costs may be important to older CLL patients, but are not captured in this analysis.

Informal caregivers

Estimates of the non-time costs of informal caregivers were informed by an American study which provided this information for caregivers of patients with advanced terminal cancer and patients receiving continuing cancer care [372]. These estimates were provided as a proportion of total informal caregiver costs. This information was used to inflate the informal care costs extracted from the EU study [371] to generate estimates of non-time costs for informal caregivers for this analysis.

Table 1: Parameters used in the societal cost calculations

Parameter	Category	Value ^a	Source
Proportion of the population in employment by age	25-49 years 50-64 years 65 years and older	80.9% 67.4% 9.5%	[360]
Rate at which the proportion of the population in employment decreases after the age of 65 years	-	20%	Ass
Short-term absenteeism in CLL patients: mean annual values per CLL patient	Sick days Physician contacts Days in hospital	7.6 14.7 8.9	[295]
Length of a working day (hours)	-	7.5	Ass
Annual presenteeism rate	-	8.45%	[367]
Proportion of time lost during presenteeism	-	100%	Ass
Annual mean hours of informal care provided per blood cancer patient	Patients dying with blood cancer Patients severely hampered by blood cancer	1,300 34	[371]
Mean out of pocket costs for patients per 28-day cycle	-	£59.64	[296]
Mean out of pocket costs for informal caregivers per 28-day cycle	Patients dying with blood cancer Patients severely hampered by blood cancer	£81.20 £2.09	[371,372]
Retirement age (years) ^b	-	65	
Median hourly earnings by age	22-29 years 30-39 years 40-49 years 50-59 years 60 years and older	£10.58 £14.37 £14.63 £13.90 £12.12	[361]
Minimum wage rate per hour	-	£6.31	[365]

Ass = assumption; CLL = chronic lymphocytic leukaemia. ^a All costs are for the year 2013; ^b Assumed to be the same for men and women.

Adapting the cost-effectiveness and cost-utility analysis to reflect a societal costing perspective

The additional costs that are relevant from a societal perspective were accrued by patients in model states as specified above, in addition to the costs that were accrued from a national health service (NHS) perspective. All unit costs are reported in British Pounds for the year 2013. All of the parameters that were used in the societal cost calculations were varied in a univariate sensitivity analysis, with parameter ranges reflecting evidence from the original data source when this was available. In the absence of such evidence, parameters were varied by 50% above and below their base case value. Results were also calculated for different starting ages for patients in the model (25 and 45, with 25 reflecting the likely youngest age at presentation [163]). Probabilistic sensitivity

analysis (PSA) was used to further consider the impact of parameter uncertainty, and the results of these simulations were used to calculate cost-effectiveness acceptability curves (CEACs).

Results

The additional discounted societal costs for the entire patient cohort across the 30 year model timeframe are presented in **Table 2** for each comparator. The costs from a NHS perspective are also presented, along with combined costs across both analytical perspectives.

The highest societal costs are associated with Comparator A (current practice in the Oxford University Hospitals NHS Trust). These costs total £4,459 per patient, with informal care costs accounting for 48% of this figure. The lowest societal costs are associated with Intervention One (genomic testing followed by first-line ibrutinib treatment in likely FCR non-responders). These costs total £3,028 per patient, with informal care costs accounting for 49% of this figure. This difference likely arises because ibrutinib treatment leads to patients spending more time in asymptomatic health states, reducing the incidence of presenteeism and the associated costs for informal carers. The breakdown of societal costs across the different cost categories is relatively constant across the five strategies. In all cases, there are no costs associated with long-term absenteeism or mortality as the starting age of patients in the model is 65 years (i.e. retirement age).

Overall, the additional societal costs for the different strategies only represent a small increase in total costs from a NHS perspective. Societal costs as a proportion of total costs range from 2.5% (Intervention One) to 5.9% (Comparator A). Consequently the cheapest (Comparator C - current practice with no genetic testing) and most expensive (Intervention One) strategies remain the same as when a NHS perspective is considered.

The results for the cost-effectiveness analysis (CEA) and CUA from a societal perspective are presented in **Table 3** with the base case results from a NHS perspective presented in **Table 4** for comparison. As the increase in mean costs per patient is small for each comparator, there is little change in the overall results. The same comparators are dominated (or extendedly dominated) in both the CEA and the CUA, and the incremental cost-effectiveness ratios (ICERs) are approximately the same, with small increases for Comparator A and small decreases for Comparator B and Intervention One. The concordance between the results of the CEA and CUA that was observed in the base case analysis is therefore replicated when this analysis is extended to a societal perspective.

Table 2: Costs for each comparator from both a societal and NHS perspective

Category	Comparator									
	A		B		C		Int 1		Int 2	
	Costs	%	Costs	%	Costs	%	Costs	%	Costs	%
Short-term absenteeism	£1,362,793	3%	£1,072,947	3%	£1,256,637	3%	£901,480	3%	£1,001,198	3%
Short-term presenteeism	£13,210,024	30%	£10,400,453	33%	£12,181,017	30%	£8,738,362	29%	£9,704,966	30%
Long-term absenteeism	£0	0%	£0	0%	£0	0%	£0	0%	£0	0%
Mortality costs	£0	0%	£0	0%	£0	0%	£0	0%	£0	0%
Informal care costs	£21,477,132	48%	£13,336,033	43%	£19,974,374	48%	£14,973,485	49%	£15,522,432	48%
OOP costs	£8,542,325	19%	£6,357,403	20%	£7,866,065	19%	£5,665,132	19%	£6,187,135	19%
Total additional societal costs (whole cohort)	£44,592,274	-	£31,166,836	-	£41,278,092	-	£30,278,458	-	£32,415,731	-
Total additional societal costs (per patient)	£4,459	-	£3,117	-	£4,128	-	£3,028	-	£3,242	-
Total costs from NHS perspective (whole cohort)	£715,759,326	-	£1,077,033,723	-	£697,044,181	-	£1,190,882,282	-	£917,896,030	-
Total costs from NHS perspective (per patient)	£71,576	-	£107,703	-	£69,704	-	£119,088	-	£91,790	-
Total costs (whole cohort)	£760,351,600	-	£1,108,200,559	-	£738,322,273	-	£1,221,160,741	-	£950,311,761	-
Total costs (per patient)	£76,035	-	£110,820	-	£73,832	-	£122,116	-	£95,031	-

Int 1 = Intervention One; Int 2 = Intervention Two; NHS = National Health Service; OOP = out of pocket.

Table 3: Results for the cost-effectiveness and cost-utility analyses from a societal perspective

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£73,832	-	-
	A	6.61	£76,035	£9,302	£9,302
	Int 2	6.65	£95,031	£545,428	EXT.DOM
	B	7.63	£110,820	£16,007	£34,062
	Int 1	7.45	£122,116	DOM	DOM
CUA	C	5.60	£73,832	-	-
	A	5.82	£76,035	£10,081	£10,081
	Int 2	5.93	£95,031	£166,523	EXT.DOM
	B	6.44	£110,820	£30,908	EXT.DOM
	Int 1	6.67	£122,116	£50,164	£54,207

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life-years; QALYs = quality-adjusted life years.

Table 4: Base case results for the cost-effectiveness and cost-utility analyses

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	C	6.37	£69,704	-	-
	A	6.61	£71,576	£7,903	£7,903
	Int 2	6.65	£91,790	£580,390	EXT.DOM
	B	7.63	£107,703	£16,133	£35,376
	Int 1	7.45	£119,088	DOM	DOM
CUA	C	5.60	£69,704	-	-
	A	5.82	£71,576	£8,565	£8,565
	Int 2	5.93	£91,790	£177,198	EXT.DOM
	B	6.44	£107,703	£31,153	EXT.DOM
	Int 1	6.67	£119,088	£50,559	£55,891

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life-years; QALYs = quality-adjusted life years.

Sensitivity analysis

As there are only small changes in the mean costs per patient for each comparator when a societal perspective is considered, the same parameter variations that change the most cost-effective strategy from Comparator A to an alternative strategy in the base case analysis are also important in this extended analysis. These results are therefore not presented. There were also few variations in the parameters that were used in the societal perspective that change the most cost-effective strategy from Comparator A, which was to be expected given the minimal contribution of societal costs to overall costs.

The only variation that did change the results was when different starting ages were considered for patients in the model. The results for the CEA and CUA from a societal perspective for patients aged 25 and 45 are presented in **Table 5** and **Table 6**. The mean costs per patient increase considerably in both cases. For 25 year old patients, the most expensive strategy is Comparator C at £707,752 per patient while the cheapest strategy is Intervention Two at £687,062 per patient. These large increases in cost arise because more CLL patients now die at a younger age, increasing the costs

associated with premature mortality. Indeed, these costs represent 99% of all additional societal costs for all five strategies. The main consequence of these cost increases is that the most cost-effective strategy in the CEA changes from Comparator A to Comparator B (current practice with ibrutinib given as refractory treatment for all patients) with an ICER of £2,943 per life-year gained. The most cost-effective strategy in the CUA also changes from Comparator A to Intervention One with an ICER of £19,933 per QALY gained. When a starting age of 45 years is considered, the mean costs per patient fall, but are still considerably higher than those incurred in the base case analysis. Comparator B is the most cost-effective strategy in both the CEA and the CUA for this patient group.

Table 5: Results for the cost-effectiveness and cost-utility analyses from a societal perspective for patients aged 25

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	Int 2	7.73	£687,062	-	-
	B	8.88	£690,438	£2,943	£2,943
	A	7.35	£697,523	DOM	DOM
	Int 1	8.73	£698,459	DOM	DOM
	C	7.00	£707,752	DOM	DOM
CUA	Int 2	6.91	£687,062	-	-
	B	7.42	£690,438	£6,701	£6,701
	A	6.47	£697,523	DOM	DOM
	Int 1	7.82	£698,459	£19,933	£19,933
	C	6.15	£707,752	DOM	DOM

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life-years; QALYs = quality-adjusted life years.

Table 6: Results for the cost-effectiveness and cost-utility analyses from a societal perspective for patients aged 45

Analysis	Comparator	Mean LYs/QALYs per patient	Mean costs per patient	ICER (excluding dominated strategies)	ICER (excluding extendedly dominated strategies)
CEA	A	7.25	£345,301	-	-
	C	6.91	£352,471	DOM	DOM
	B	8.70	£354,766	£6,530	£6,530
	Int 2	7.57	£365,165	DOM	DOM
	Int 1	8.54	£370,966	DOM	DOM
CUA	A	6.38	£345,301	-	-
	C	6.07	£352,471	DOM	DOM
	B	7.27	£354,766	£10,616	£10,616
	Int 2	6.77	£365,165	DOM	DOM
	Int 1	7.65	£370,966	£43,272	£43,272

CEA = cost-effectiveness analysis; CUA = cost-utility analysis; DOM = dominated; EXT.DOM = extendedly dominated; ICER = incremental cost-effectiveness ratio; Int 1 = Intervention One; Int 2 = Intervention Two; LYs = life-years; QALYs = quality-adjusted life years.

These results indicate that at younger ages, a strategy which includes some form of ibrutinib treatment is likely to be the most cost-effective option from a societal perspective. This is because fewer patients die when undergoing ibrutinib treatment, which reduces mortality costs compared to strategies in which patients don't receive ibrutinib. In addition, these results indicate that discordance between the results of the two forms of extra-welfarist economic evaluation is more

likely at younger ages. For 65 year old patients, the most cost-effective strategy in both the CEA and the CUA is Comparator A. This changes to Comparator B for 45 year old patients. However, for 25 year old patients the most cost-effective strategy differs between the CEA and CUA.

A final point to note in relation to patient age is that the societal analysis did not include the costs associated with admission to a nursing home. The use of ibrutinib treatment in older patients may prevent such admissions as patients with effectively treated disease are better able to cope with other comorbidities. However, this effect is not captured in the model. If it were captured, it is possible that strategies which include ibrutinib treatment would also be cost-effective in older patients when a societal perspective is considered.

Probabilistic sensitivity analysis

The results of the PSA allowed CEACs to be calculated (**Figures 1 and 2**) which are similar to those for the analysis which did not consider societal costs. These CEACs confirm that the most cost-effective strategy at a threshold of £30,000 per life-year/QALY gained is Comparator A, as per the analysis from a NHS perspective. Again, the CEACs for the CUA suggest that Intervention One becomes the most cost-effective strategy at a threshold of approximately £50,000.

Figure 1: Cost-effectiveness acceptability curves for the cost-effectiveness analysis from a societal perspective

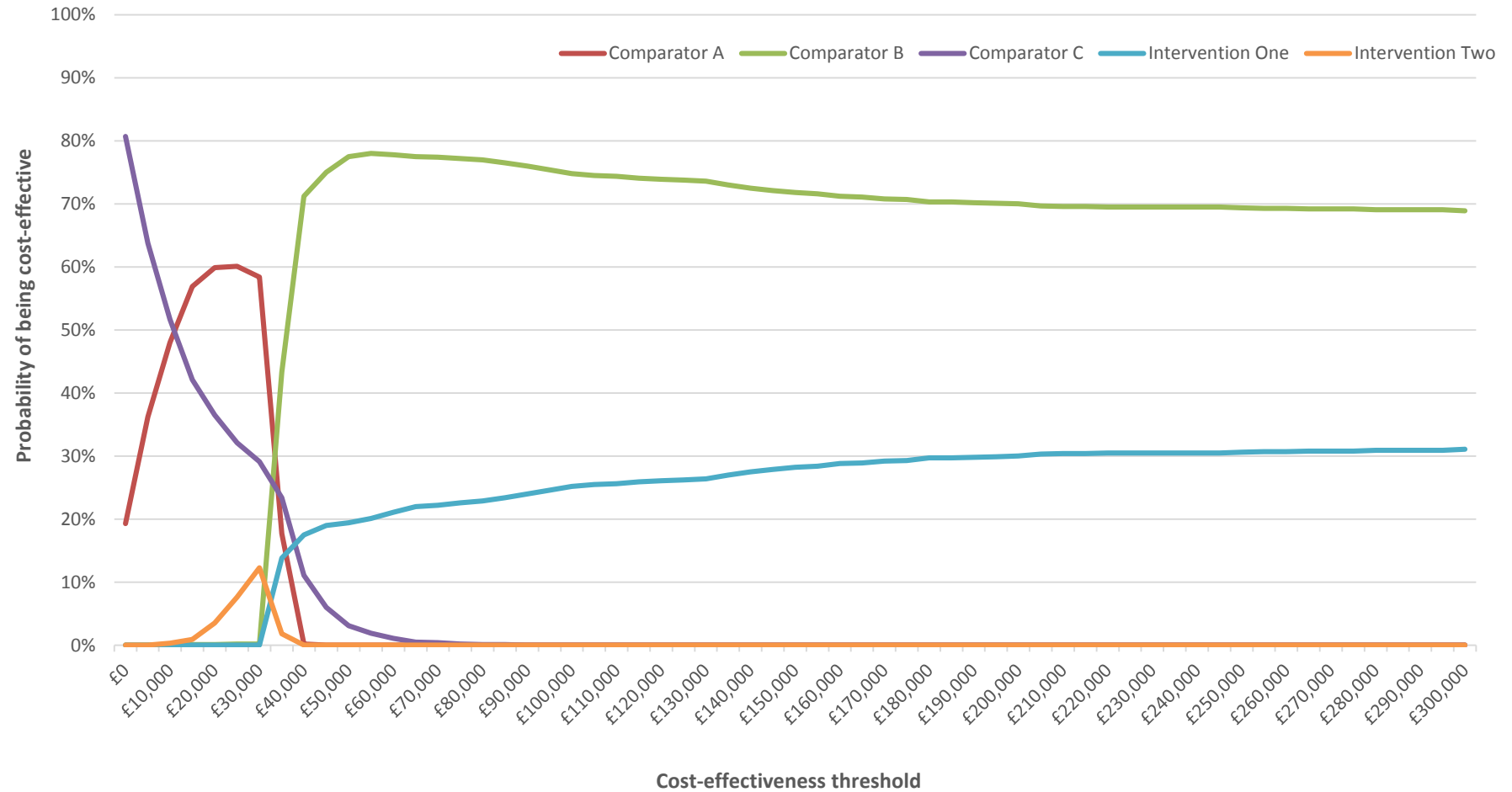
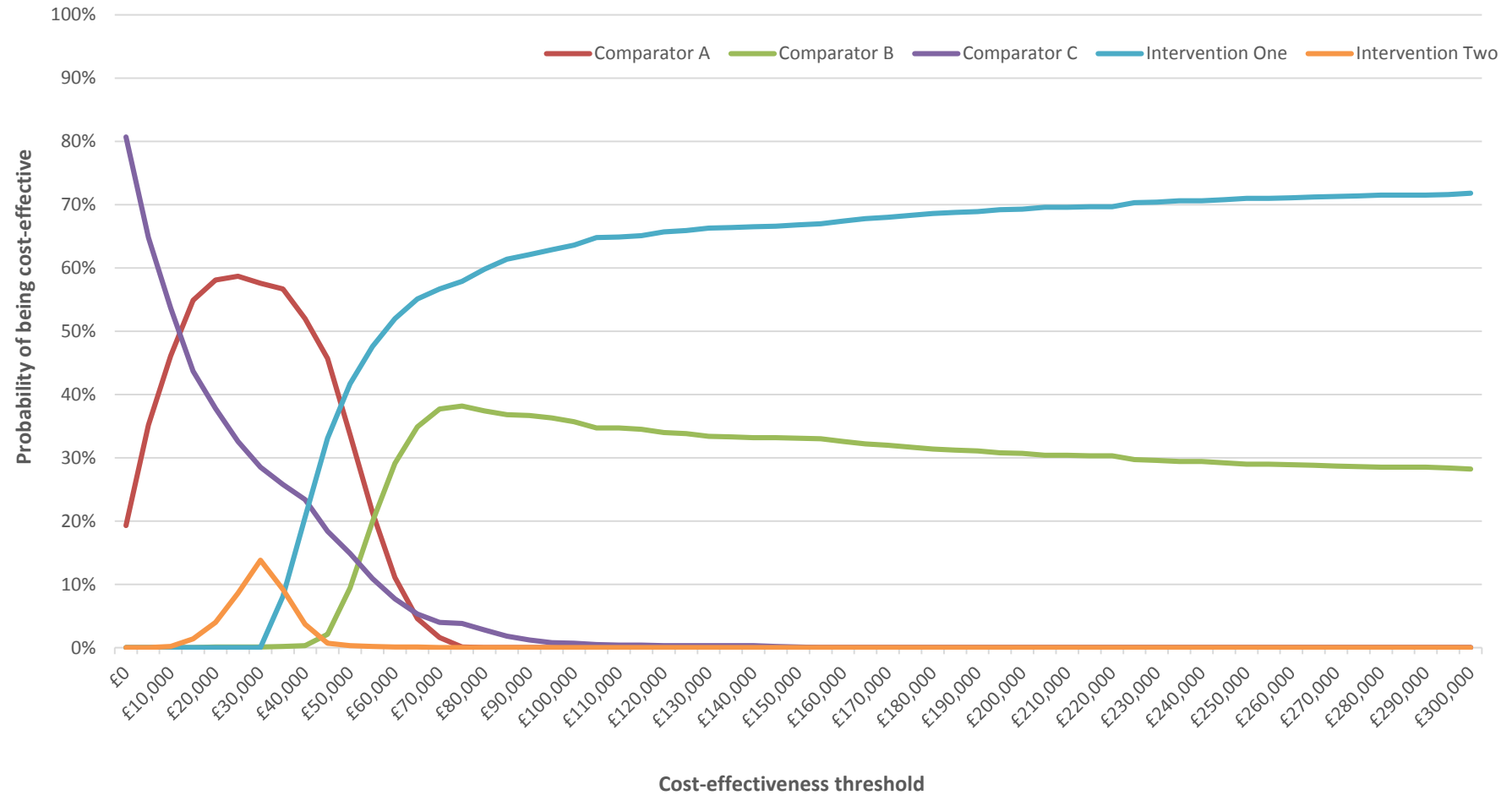


Figure 2: Cost-effectiveness acceptability curves for the cost-utility analysis from a societal perspective



Appendix 37: Expected value of perfect information results presented graphically

Figure 1: Expected value of perfect information from a NHS perspective (cost-effectiveness analysis)

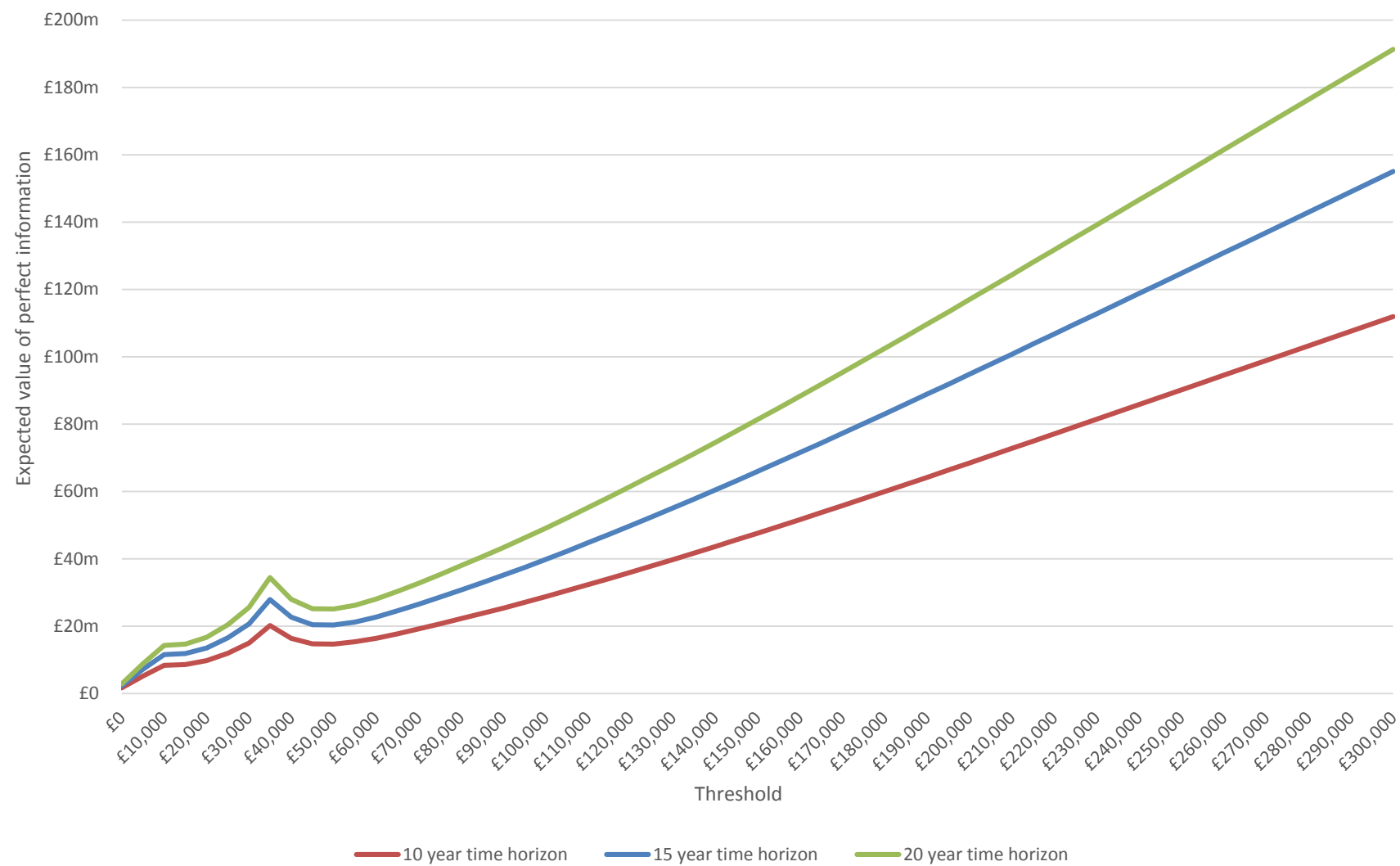


Figure 2: Expected value of perfect information from a NHS perspective (cost-utility analysis)

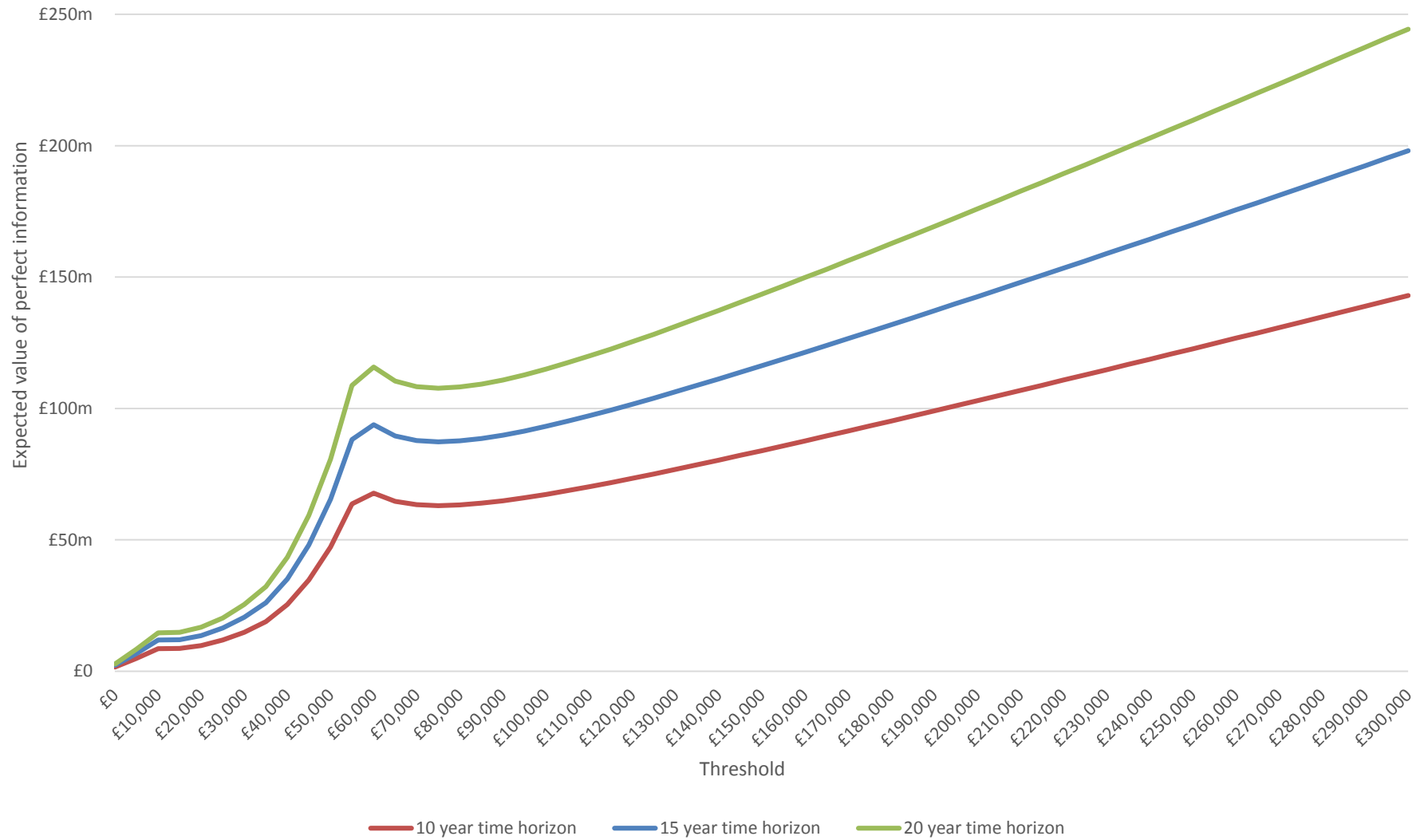


Figure 3: Expected value of perfect information from a societal perspective (cost-effectiveness analysis)

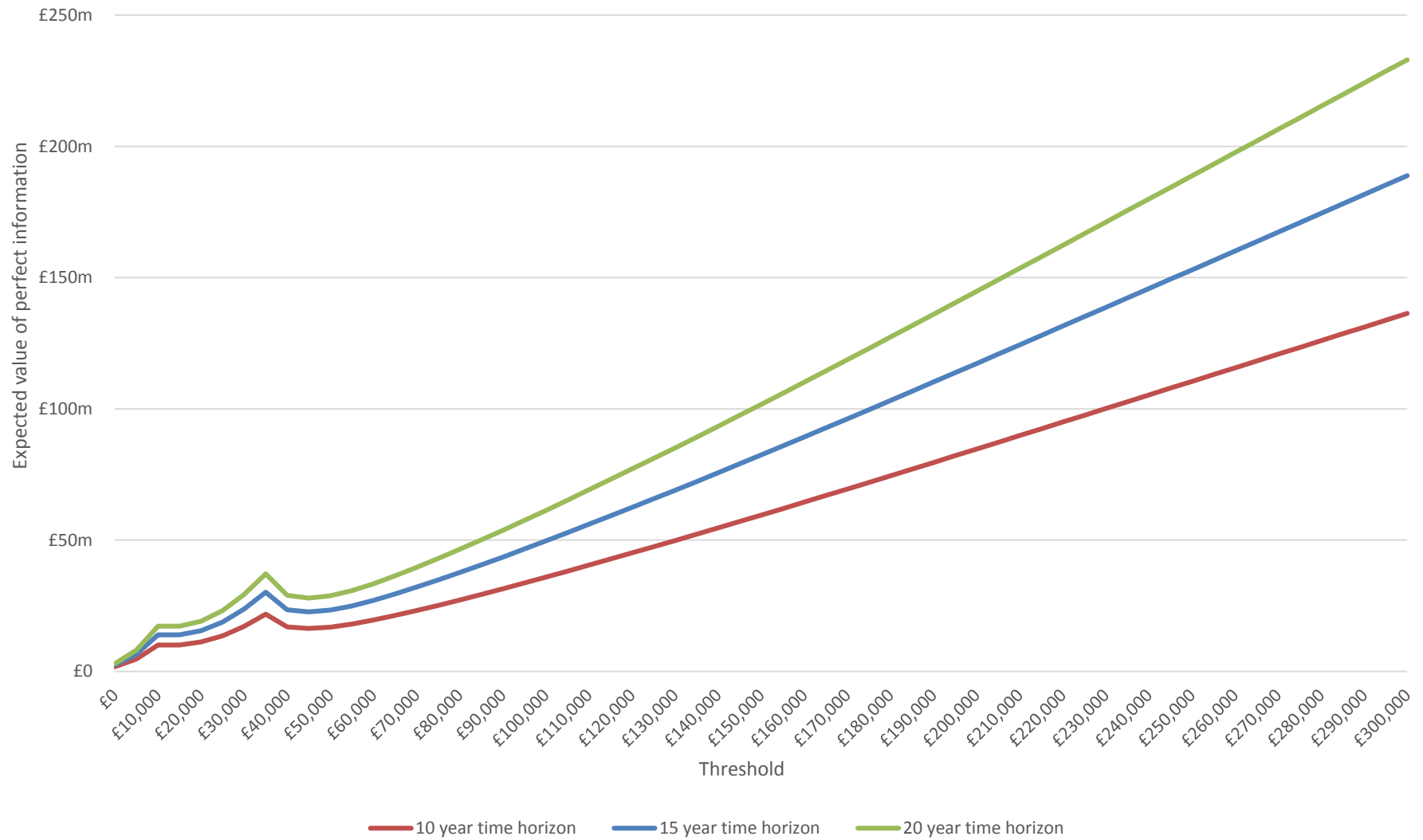
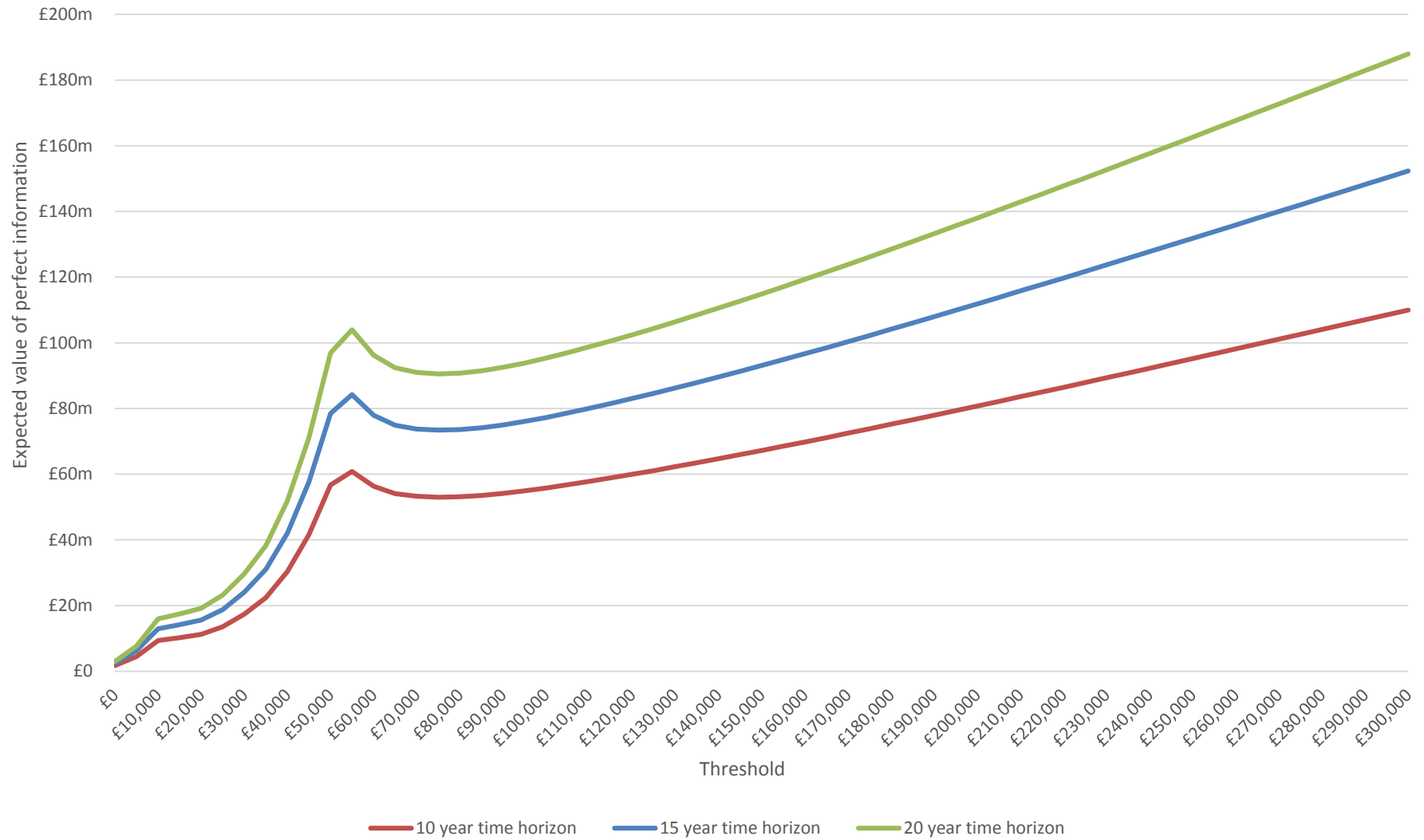


Figure 4: Expected value of perfect information from a societal perspective (cost-utility analysis)



Appendix 38: Survey of healthcare decision-makers – methods and results

This appendix describes the various steps that were taken to design and conduct the survey of healthcare decision-makers. The results of this survey are also presented.

Methods

The main aim of the survey was to elicit opinions on which economic evaluation results respondents believed captured the true costs and benefits of genomic testing in chronic lymphocytic leukaemia (CLL). Secondary aims were to establish a consensus regarding an adoption decision in this context, and to ask respondents to consider whether standard health technology assessment (HTA) methods should be adapted to reflect the challenges of conducting economic evaluations of genetic and genomic tests.

Target respondents fell into three categories:

1. Health economists who hold policy-making or commissioning positions. Such respondents could include individuals working for HTA organisations such as the National Institute for Health and Care Excellence, the Canadian Agency for Drugs and Technologies in Health, the Pharmaceutical Benefits Advisory Committee in Australia and equivalent organisations in other countries, and also local healthcare commissioners in the UK.
2. Health economists who either work in the area of genomics, have published papers on the topic of welfarism versus extra-welfarism, or have a background in cost-benefit analysis (CBA).
3. Individuals with relevant backgrounds in two or more of genomics, policymaking, science research and medicine.

Respondents were identified in several ways. First, respondents were contacted who appeared on existing lists of health economists working in the areas of health economics and genomics (e.g. the [Health economics and genomics](#) blog). A number of relevant health economists were also identified by reviewing the author lists of publications in this field, in particular authors of the publications included in the two literature reviews presented in Chapter Two. Health economists who hold policy-making or commissioning positions were identified via HTA agency websites. Other relevant respondents were either already known to the author, or were identified by asking existing respondents to recommend additional respondents who would meet the recruitment criteria. 91 potential respondents were targeted in total. As the recruitment process was not systematic, it is possible that some relevant respondents were not asked to participate, or that those who were included were not representative of the entire potential pool of respondents. However, there are few health economists working in this area and almost all are known to the author, so comprehensive coverage of this sub-group of respondents was anticipated.

The survey design process had two stages. First, a document was created which provided a summary of the welfarist and extra-welfarist economic evaluation results. Information was also provided on the various analytical extensions described in Chapter Six. This document was three pages in length. A peer-reviewed publication reporting the full methods and results of these economic evaluations would likely be much longer, but a balance was sought between providing respondents with enough information to be able to provide informed responses to the survey questions, whilst at the same time encouraging participation. A copy of this document is provided in **Appendix 39**.

The second stage was to develop a series of questions based on these materials which reflected the main aims of this survey. A provisional list of questions was assembled with the assistance of colleagues working on this study, and this was then piloted (alongside the supporting materials) in a convenience sample of nine health economists working at the Health Economics Research Centre at the University of Oxford.

The final list of questions is provided in **Appendix 40**. Respondents were initially presented with some introductory text before indicating their prior experience of conducting economic evaluations. Respondents were then asked a series of questions relating to the supporting materials. First, they were asked to indicate which economic evaluation approaches best captured the costs and benefits of genomic testing in CLL, before recommending one of four adoption decisions. Then they were asked whether CBA results would be useful for healthcare decision-makers both in this context, and for other healthcare interventions. Finally, they were asked whether they thought standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests. A journal article was provided as a reference point for this question to ensure that all respondents were aware of the key challenges in this context. Respondents were asked to justify their responses to all questions. The wording of some of these open-ended questions was tailored to reflect the answers that respondents had previously provided in the survey. In addition to the information collected in the survey, respondent location and occupation was also captured.

The survey was conducted in July and August 2015 using the LimeSurvey online survey application [194]. Screenshots are provided in **Appendix 40**. Potential respondents were contacted by email to request participation. A copy of the supporting materials was attached to the email in pdf format, and respondents were directed to the online survey at the following web address: <http://limesurvey.msd.ox.ac.uk/index.php/441274/lang-en>. Respondents were asked to complete the survey within one month. After this date non-responders were sent an email reminder and asked to complete the survey within two weeks. Only those respondents who had been contacted to request participation were permitted to respond to the survey (anonymous responses were not permitted).

Following the survey end date data were extracted from LimeSurvey and analysed in Microsoft Excel and Stata 12 [240,277]. Responses were evaluated for the whole group and also potentially relevant sub-groups based on responder occupation, location, and prior experience of conducting economic evaluations. All responses to open-ended questions were reviewed to identify common themes.

After the survey had been conducted an error was discovered in the WTP calculations underlying the CBA. Specifically, the levels chosen for the test effectiveness attribute were discovered to be incorrect. The CBA results presented to survey respondents were therefore not the same as the final results reported in this thesis. In both cases the net benefit for genetic testing was negative and the net benefit difference between the genetic and genomic tests was positive. The main changes were that willingness-to-pay (WTP) for both tests became negative, and the net benefit associated with genomic testing also became negative. The net benefit confidence intervals for both tests also no longer spanned zero. The impact of this error on the validity of the survey results is discussed in the results section.

Results

A complete list of survey responses is provided in **Appendix 41**. The following three sections describe the characteristics of respondents, provide an overview of respondent decisions, and outline how these decisions were justified.

Respondent characteristics

Thirty-eight responses were received (response rate: 42%). The characteristics of survey respondents are described in **Table 1**. Most (37%) were from the UK and 79% of all respondents were from either the UK, Canada, USA or Australia. Most respondents (82%) were health economists. All respondents who were not health economists were from the UK. Most respondents (71%) had been lead health economists on one or more cost-effectiveness (CEAs) or cost-utility analyses (CUAs). Only 21% of respondents had been lead health economists on one or more CBAs, with most (61%) having limited or no experience of welfarist approaches.

Table 1: Respondent characteristics (N=37)

Variable	n (%)
Respondent location:	
Australia	7 (18%)
Canada	4 (11%)
Germany	1 (3%)
Ireland	2 (5%)
Netherlands	2 (5%)
New Zealand	2 (5%)
Norway	1 (3%)
UK	14 (37%)
USA	5 (13%)
Number of respondents who were health economists	31 (82%)
Knowledge of CEA/CUA:	
I have been the lead health economist on one or more CEAs or CUAs	27 (71%)
I have been involved in one or more CEAs or CUAs, but not as the lead health economist	3 (8%)
I have limited experience of conducting CEAs or CUAs	5 (13%)
I have no experience of conducting CEAs or CUAs	3 (8%)
Knowledge of CBA:	
I have been the lead health economist on one or more CBAs	8 (21%)
I have been involved in one or more CBAs, but not as the lead health economist	7 (18%)
I have limited experience of conducting CBAs	17 (45%)
I have no experience of conducting CBAs	6 (16%)

CBA = cost-benefit analysis; CEA = cost-effectiveness analysis, CUA = cost-utility analysis.

Overview of respondent decisions

Respondents were asked to make four decisions in the survey. An overview of their responses is presented in **Tables 2-5**, and a summary is provided below. The number of respondents in some of the sub-groups is small, hence these results should be interpreted with caution.

Which approach(es) best capture the costs and benefits of genomic testing in CLL?

Most respondents (39%) selected CUA from a societal perspective, followed by CUA from a national health service (NHS) perspective (34%). Health economists were more likely to select CUA from a societal perspective than respondents who were not health economists (42% vs. 29%). Although a

societal perspective was favoured for the CUA compared to a NHS perspective (39% vs. 34%), the reverse was true for the CEA (24% vs. 18%). CBA was the least selected approach (13%). Around a quarter of respondents (26%) said that they did not have enough information to make a decision: health economists were more likely to provide this response than respondents who were not health economists (32% vs. 0%).

Some differences were observed by respondent location. UK respondents were less likely to select an approach which adopts a societal perspective, while Australian respondents found CUAs to be much more favourable, whatever the perspective. Canadian respondents were more likely to say that they had not been provided with enough information, while American respondents were more likely to say 'none of the above' and more likely to choose CUA.

The results were also broken down by prior respondent knowledge of economic evaluation. These results indicated that respondents who had been the lead health economist on one or more CBAs were more likely to select CEA (from either a NHS or societal perspective), CUA (from a NHS perspective) or indicate that they had not been provided with enough information. However, they were less likely to select CUA (societal perspective).

What adoption decision should be made regarding genomic testing in CLL?

Most respondents (45%) said that they would delay a decision, while 34% said that they would either adopt or adopt with conditions. Respondents who were not health economists were more likely to adopt than health economists (43% vs. 10%). UK respondents were very keen to adopt, whereas Australians preferred to delay. Canadians and Americans were even more keen to delay or not adopt. Respondents who had been the lead health economist on one or more CEAs or CUAs were generally more cautious in their recommended adoption decisions, with 52% recommending delaying a decision. Those who had been the lead health economist on one or more CBAs were very cautious, with nobody recommending outright adoption and 63% recommending delaying a decision.

Would it be helpful for healthcare decision-makers at the national level to consider CBA results alongside CEA or CUA results when determining whether to adopt genomic testing for CLL patients?

Respondents were generally uncertain about whether this would be useful, with yes/no votes split approximately equally. Respondents who were not health economists were in greater agreement (43% vs. 23%), whereas more health economists said no than yes (32% vs. 23%). Australian and American respondents were more likely to disagree with this statement, while Canadian respondents were more likely to agree. UK respondents were more evenly split. Those who had been the lead health economist on one or more cost-benefit analyses were also more likely to be in agreement with this statement.

Should standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests?

Most respondents replied 'yes', or 'yes in specific circumstances' (71%). There was little difference in the results by whether the respondent was a health economist. UK respondents were less likely to respond 'yes' whereas Australian and American respondents were more likely to say 'yes'. Canadian respondents followed the overall results. There were no differences between respondents with different experiences of CEA, CUA and CBA.

Table 2: Which approach(es) best capture the costs and benefits of genomic testing in chronic lymphocytic leukaemia?

Approach ^a	Overall	Health economist?		Respondent location				I have been the lead health economist on one or more...	
		Yes (n=31)	No (n=7)	UK (n=14)	Australia (n=7)	USA (n=5)	Canada (n=4)	...CEAs/CUAs (n=27)	...CBAs (n=8)
CEA (NHS)	24%	23%	29%	21%	29%	20%	25%	22%	38%
CEA (Societal)	18%	19%	14%	7%	29%	0%	25%	19%	38%
CUA (NHS)	34%	35%	29%	36%	57%	0%	25%	33%	50%
CUA (Societal)	39%	42%	29%	21%	57%	60%	25%	41%	25%
CBA	13%	13%	14%	14%	14%	0%	25%	15%	13%
Lack of info ^b	26%	32%	0%	21%	14%	20%	75%	33%	38%
None ^c	13%	13%	14%	7%	14%	40%	25%	15%	13%

CBA = cost-benefit analysis; CEA = cost-effectiveness analysis; CUA = cost-utility analysis, NHS = National Health Service perspective. ^a Respondents could select more than one approach, hence these percentages do not sum to 100%. ^b The exact wording for this answer was: "I have not been provided with enough information to answer this question". ^c The exact wording for this answer was: "I don't believe that any of these results capture the true costs and benefits of genomic testing in CLL".

Table 3: What adoption decision would you recommend?

Decision ^a	Overall	Health economist?		Respondent location				I have been the lead health economist on one or more...	
		Yes (n=31)	No (n=7)	UK (n=14)	Australia (n=7)	USA (n=5)	Canada (n=4)	...CEAs/CUAs (n=27)	...CBAs (n=8)
Adopt	16%	10%	43%	36%	0%	0%	0%	7%	0%
Adopt with conditions	18%	16%	29%	21%	14%	0%	0%	15%	13%
Delay decision	45%	48%	29%	21%	57%	80%	75%	52%	63%
Do not adopt	21%	26%	0%	21%	29%	20%	25%	26%	25%

CBA = cost-benefit analysis; CEA = cost-effectiveness analysis; CUA = cost-utility analysis. ^a The exact wording for the four possible responses to this question was: 1. Adopt genomic testing; 2. Adopt genomic testing, but with some conditions e.g. for specific patient subgroups, only after later lines of treatment; 3. Delay a decision until additional research is undertaken, and 4. Do not adopt genomic testing.

Table 4: Would it be helpful for healthcare decision-makers at the national level to consider CBA results alongside CEA or CUA results when determining whether to adopt genomic testing for chronic lymphocytic leukaemia patients?

Response ^a	Overall	Health economist?		Respondent location				I have been the lead health economist on one or more...	
		Yes (n=31)	No (n=7)	UK (n=14)	Australia (n=7)	USA (n=5)	Canada (n=4)	...CEAs/CUAs (n=27)	...CBAs (n=8)
Yes	26%	23%	43%	29%	14%	20%	50%	22%	38%
No	29%	32%	14%	21%	43%	40%	25%	33%	13%
Unsure	45%	45%	43%	50%	43%	40%	25%	44%	50%

CBA = cost-benefit analysis; CEA = cost-effectiveness analysis; CUA = cost-utility analysis.

Table 5: Do you think that standard health technology assessment methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests?

Response ^a	Overall	Health economist?		Respondent location				I have been the lead health economist on one or more...	
		Yes (n=31)	No (n=7)	UK (n=14)	Australia (n=7)	USA (n=5)	Canada (n=4)	...CEAs/CUAs (n=27)	...CBAs (n=8)
Yes	55%	55%	57%	43%	71%	60%	50%	56%	63%
Only in specific circumstances	16%	19%	0%	7%	14%	20%	25%	19%	13%
No	29%	26%	43%	50%	14%	20%	25%	26%	25%

CBA = cost-benefit analysis; CEA = cost-effectiveness analysis; CUA = cost-utility analysis; HTA = health technology assessment.

Justification for respondent decisions

Respondents were asked a number of open-ended questions related to the decisions that they made in the survey. These responses are provided in full in **Appendix 41**, and common themes are summarised here. Respondent identification numbers are provided in parentheses.

Which approach best captures the costs and benefits of genomic testing in CLL?

Following this question, respondents were asked to explain why their selected approaches were chosen. Respondents were able to select more than one approach, so several different combinations of approaches were ultimately selected. These are broadly grouped as follows:

- Cost-effectiveness analysis (NHS or societal perspective): One respondent noted that linking genomic data to health outcomes seemed to be the most relevant approach (s36), while another commented that the results from a societal perspective were useful in order to make an effective budget bid for additional resources from the governing body (s3). A third respondent noted that measures of effectiveness are more robust and easier to defend than utility or monetary measures (s10).
- Cost-utility analysis (NHS perspective): This was the second most popular choice. A variety of reasons were given. Some noted that CEAs do not incorporate quality of life benefits (s27, s38), that a societal perspective introduces too many uncertainties (s27) and that CBAs are difficult to interpret (s27). Others focused on the role of the healthcare decision-maker, stating that this was the best approach to prioritise investments from a fixed health budget (s3), and that a generic measure should be used (i.e. quality-adjusted life years (QALYs)) if the opportunity costs fall on health (s5, s6).
- Cost-utility analysis (societal perspective): This was the most popular choice. One respondent selected this approach as CBAs do not include treatment effects and CEAs do not include quality effects (s4). Others noted that this was the most transparent and comprehensive approach (s21, s23, s24, s30) and that consensus is most likely for this approach (s9).
- Cost-benefit analysis: This was the least selected approach. Those that favoured CBA believed that it best captured the costs and benefits of genomic testing (although it may not necessarily be the best approach for informing decision-making) (s13). It was also noted that it was important to understand the choices that people would make in practice, and why (s18), and that given the high cost of ibrutinib, public engagement surrounding its coverage may be warranted (s29).
- Combinations of approaches: Various combinations of approaches were selected. In most cases, respondents emphasised that multiple sets of results would strengthen decision-making. For example, a respondent who selected both the NHS and societal versions of the CEA and CUA noted that in combination, these approaches would allow a reasonable characterisation of the intervention benefits (s11). Respondents who selected CEA and CUA from a NHS perspective justified this decision by noting that these approaches represented standard methods (s13) and were most familiar and 'concrete' (s19, s32). Two respondents selected every approach on the basis that each analysis provides slightly different (but still useful) information (s17, s29).
- Lack of information: Those respondents who were unable to make an adoption decision identified a number of pieces of information that would be required to facilitate decision-

making. Several respondents stated that they needed more information on the objectives of the decision-maker that the analysis has been undertaken to inform, and on the budget that the decision-maker was allocating (s1, s2, s14, s20, s25, s34). Further information was also required on the exact nature of the genomic test (s22) and the choices made in the discrete choice experiment (DCE) (s33).

- None of the options: Finally, 13% of respondents did not believe that any of the results captured the true costs and benefits of genomic testing in CLL. The reasons given for this decision included a lack of confidence in the CBA (as WTP seems too low) and in the CEA and CUA (as appropriate strategies were not modelled) (s7). Others required a budget impact analysis (s8) or suggested that a better starting point might be a cost-consequence analysis (CCA) (s35).

What adoption decision would you recommend?

Following this question, respondents were asked to explain why they would make this decision. These responses are grouped by adoption decision as follows:

- Adopt: Only 16% of respondents recommended unconditional adoption. The reasons cited included that the incremental cost-effectiveness ratio versus current practice is favourable (s6, s10, s36), that it is important to avoid the side-effects associated with chemotherapy where possible (s31) and that the CBA returns a positive net benefit for genomic testing (s33).
- Adopt with conditions: The conditions mentioned by respondents in this category included adoption alongside prospective collection of better data (s4, s18, s22), adoption with managed entry (s9) and adoption in certain sub-groups (s14, s16).
- Delay: 45% of respondents recommended delaying a decision. The most common reason given was that insufficient information had been provided in the supporting materials (s1, s2, s7, s8, s20, s21, s25, s26, s29, s37). Others would delay until the price of ibrutinib falls (s15, s21), until further information is available on genomic testing in subgroups (s27) or until information is available separately on the performance of genomic testing and ibrutinib treatment (s32, s34).
- Do not adopt: One in five respondents recommended that genomic testing should not be adopted. Most of these respondents noted that comparators which included genetic testing offered better value for money (s11, s12, s23, s24, s28, s30, s35). Another respondent cited the limited evidence on clinical benefit (s13) while a further respondent noted that the estimate of expected value of perfect information suggested that a future trial would be unlikely to improve the evidence base in a cost-effective manner (s23).

Please explain why decision-makers should consider CBA results when deciding whether to adopt genomic testing in CLL

A variety of reasons were provided. Respondents noted the importance of considering the national perspective (s7) and of considering evidence on real decision-making (s18). Some noted that the value of genomic testing is masked by the cost of treatment in CEAs and CUAs, but not in CBAs (s22). Two respondents commented that CEAs and CUAs do not incorporate the value of knowing (s29, s38), while a further respondent noted that CBAs may be more readily calculated (s36).

Please explain why decision-makers should not consider CBA results when deciding whether to adopt genomic testing in CLL

Respondents who thought that decision-makers should not consider CBA results in this context, or who were unsure, were grouped together in this category. Some focused on the advantages of the status quo, noting that most of the time CUAs (if done well) are sufficient (s3) and already capture patient preferences via utility weights (s27). Others stated that the costs of genomic testing fall on NHS budgets, so this is where the opportunity cost lies (s6, s12, s23). Another group of respondents focused on perceived deficiencies in the welfarist approach, stating that the additional costs and benefits of undertaking CBA were unclear (s2, s28, s32) and that the methods for determining WTP are not sufficiently reliable (s9, s21, s24, s34).

A final group of respondents considered the perspective of the decision-maker. It was noted that CBAs introduce uncertainty over interpretation and which result is believed (s14, s17, s19, s20) and may over-complicate the process (s15, s23). CBAs also place too much responsibility on the decision-maker to select an appropriate evaluative framework (s14). Furthermore, the welfarist and extra-welfarist approaches are alternative paradigms with overlaps, and double-counting is possible (s35). One respondent suggested that it would be more helpful to count non-health benefits and harms than attempt to place a value on them (s11).

A small number of respondents considered issues surrounding the use of CBAs at the local level. Concerns were expressed that local decision-makers would be more focused on budgetary impact (s7, s21), that information on real decision-making is not always useful at this level (s17, s18) and that CBAs may not be appropriate or even understood at this level of decision-making (s4, s13, s33).

Are there any circumstances in which decision-makers should consider CBA results alongside CEA or CUA results when deciding whether or not an intervention should be adopted in a healthcare system?

Respondents were next asked to take a broader view of the use of CBAs in healthcare. There were mixed views regarding the wider application of this approach. Respondents who were in favour of wider application primarily focused on the importance of capturing information on patient preferences (s8) and non-health factors (s21, s22, s26, s29, s38). Particular decision contexts where CBAs would likely be useful were also highlighted, including:

- Evaluation of rare diseases (s31);
- Clinical guidelines (s23);
- When tests have different attributes but similar health outcomes (s27, s36);
- To inform patient co-pay or out of pocket cost policies (s37);
- When bidding for additional resources (s3);
- For investment choices across sectors (s17, s28);
- When considering alternative mode of delivery interventions (s15) or semi-elective treatments e.g. genetic counselling (s19) or complex technologies (s33).

Two respondents expressed neutral viewpoints, noting that genomic technologies were not unique, so if welfarist approaches are considered in genomics, they should be considered for all interventions (s5, s25). The remaining respondents had a variety of concerns. It was noted that WTP values are not always appropriate for a public health system (s2), that methods are not good enough, and that there is no political will (s9). Some were generally not in favour, but described scenarios in which welfarist approaches might become valid such as if the remit of a healthcare system is to maximise welfare (s12), or if empirical work is undertaken which informs decision-

makers how to use both sets of evidence (s14). Two further respondents noted that it was more helpful to see disaggregated 'other' outcomes rather than CBAs i.e. cost-consequence analyses (s11, s35).

How would you like to see standard HTA methods adapted in genomics?

The wording of this question was tailored to reflect the answers that respondents had previously provided. Those respondents who thought that standard HTA methods should not be adapted were asked a different question (see below). Overall, 55% of respondents were in agreement that standard HTA methods should be adapted. However, most respondents focused on why methods should be adapted rather than suggesting solutions to the problems that were identified.

Key challenges that were cited included the fact that QALYs are not appropriate in all circumstances (s11, s28, s36), that there were issues surrounding the evaluation of multiple comparators (s21), and that genomic testing cannot be addressed in a HTA as if it is a diagnostic test as the costs and outcomes associated with downstream interventions also need to be considered (s27). It was also noted that the costs of 'finding' patients should be included in HTAs (s4) and that current methods are not fit for purpose with respect to the collection and analysis of genomic effectiveness data (s15, s24). Finally, several respondents commented that the value attached to personal utility should be explored, with consideration given to how this can be captured in an economic evaluation (s19, s21, s31).

General comments regarding the next steps were made by several respondents. One noted that techniques should evolve as they encounter particular challenges, not atrophy because of a failure to adapt (s10). Another stated that the first step should be to decide which technologies cannot be appraised using current HTA methods, which non-health outcomes are important, and how these should be captured. Only then should methods change (s14). A third respondent noted that decision-makers should reconsider the evidence that is required to make decisions when they matter (and not years afterwards) (s21), and that decision-makers also need to be willing to make decisions under conditions of greater uncertainty (possibly with the assistance of value of information analysis) (s37). Finally, the more widespread use of the capabilities approach was also suggested as a possible way forward (s28), along with the use of DCEs to capture patient preferences and heterogeneity (s17, s21, s29).

Why would you not want to see standard HTA methods adapted in genomics?

Only respondents who thought that standard HTA methods should not be adapted (or should only be adapted in certain circumstances) were asked this question. Although 45% of all respondents fell into this category, only a small number of common themes emerged. A number of respondents noted that if methods are adapted for genomics, they should be adapted for all interventions (s5, s9, s13, s22, s30, s31). Others emphasised that the only issues related to data availability, not methods (s23) and any necessary other factors could be considered in committee (s6, s12). Finally, some respondents noted that there is no consensus on the need to capture non-health outcomes (s2), and that doing so may worsen population health but improve quality of life in some other way, outside the remit of the budget that decision-makers were originally meant to allocate (s2, s12).

Impact of the calculation error on the survey results

It was previously noted that the supporting materials presented to respondents contained an error in the calculations underlying the CBA results. The main consequence of this error was that participants were presented with WTP estimates for genetic and genomic testing that were much

higher than the true WTP estimates. The most likely way in which the results of the survey would have changed if respondents had been presented with the correct WTP estimates is that respondents would have been less likely to value the CBA results, and it would have been less likely that these results would have informed their adoption decisions. However, only 13% of respondents said that the CBA results best captured the costs and benefits of genomic testing in CLL, and of these respondents, nobody cited the CBA results (i.e. the positive WTP values) as being a factor in their recommended adoption decision. Only one respondent (s33) suggested adoption on the basis of the positive net benefit in the CBA. Hence, it was concluded that this error was unlikely to have meaningfully affected the responses provided, and thus the interpretation of the survey results presented in this thesis remains valid.

Appendix 39: Supporting materials provided to survey respondents

Is genomic testing cost-effective? A case study in chronic lymphocytic leukaemia

James Buchanan¹, Sarah Wordsworth¹, Ruth Clifford^{2,3}, Pauline Robbe², Anna Schuh^{4,5}, Samantha JL Knight^{6,7}

¹ Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK; ² The Oxford Biomedical Research Centre Haemato-Molecular Diagnostic Service, John Radcliffe Hospital, Oxford, UK; ³ Oxford Cancer & Haematology Centre, Churchill Hospital, Oxford, UK; ⁴ BRC/NHS Oxford Molecular Diagnostics Centre, Oxford University Hospitals, Oxford, UK; ⁵ Department of Oncology, University of Oxford, Oxford, UK; ⁶ NIHR Biomedical Research Centre, Oxford, UK; ⁷ Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.

BACKGROUND

Genetic tests are currently used in healthcare systems around the world to diagnose inherited single gene disorders or generate information on patient disease status. New genomic technologies such as targeted next generation sequencing (NGS) can yield information on multiple genetic changes, allowing disease management to be stratified and treatments to be targeted at those likely to benefit, reducing patient morbidity and treatment costs. However, these technologies have not been implemented widely, in part due to a lack of evidence to support their use. Economic evidence could inform the translation of these tests into clinical practice, but there are several methodological challenges, including selecting the appropriate type of economic evaluation. It is this challenge that is the subject of this survey.

Chronic lymphocytic leukaemia (CLL) and genomic testing

CLL is the most common adult leukaemia in the Western world (annual incidence 4 cases per 100,000 population). As many patients are asymptomatic on presentation and CLL often initially follows an indolent clinical course, chemotherapy is usually only offered to symptomatic patients. First-line therapy for patients who can tolerate aggressive treatments is usually fludarabine, cyclophosphamide and rituximab (FCR) combination chemotherapy, however, 25% of patients will not maintain a response. Genetic changes (non-inherited mutations) contribute to this lack of response, and some guidelines recommend the use of genetic testing to guide treatment selection. However, the tests that are currently used (e.g. Sanger sequencing) generally only consider single genes and cannot provide detailed genetic information. Consequently, only a third of non-responders are currently identified.

New genomic testing approaches such as targeted NGS could improve treatment selection. This is a more automated and higher throughput form of sequencing, offering a wider view of a patient's genome at increased resolution. By providing more detailed information about multiple genetic changes, targeted NGS could reduce unnecessary treatment and associated side-effects, identifying two-thirds of FCR non-responders. However, it is yet to be translated into clinical practice, and there is little information available on the costs and benefits of testing. In parallel to the development of genomic testing in CLL, a new treatment called ibrutinib is also being developed which could replace FCR treatment for some patients, improving survival and quality of life. Ibrutinib is expensive, but has recently been approved by the UK Cancer Drugs Fund.

Decision problem

An economic evaluation would provide useful information to decision-makers tasked with deciding whether to implement genomic testing in CLL. Many healthcare decision-makers around the world (e.g. NICE, CADTH, PBAC) require that economic evaluations take an extra-welfarist approach. This usually takes the form of either cost-effectiveness analysis (CEA: health outcomes measured in life years) or cost-utility analysis (CUA: health outcomes measured in QALYs). In both cases, a healthcare provider perspective is often adopted. However, as the benefits of genomic testing in CLL are both clinical and process-related (e.g. shorter time to test result), alternative economic evaluation approaches (e.g. cost-benefit analysis (CBA): health outcomes measured in monetary terms) might be useful. It may also be informative to consider a wider societal perspective (e.g. including productivity costs). We therefore carried out an economic evaluation of genomic testing for CLL patients in a UK setting using three approaches (CEA, CUA, CBA) in order to consider the most appropriate approach in this context. The CEA and CUA were conducted from both a healthcare provider and societal perspective. The CBA was informed by a discrete choice experiment (DCE).

METHODS

Cost-effectiveness analysis (CEA) and cost-utility analysis (CUA)

A Markov model was constructed to conduct the CEA and CUA, informed by current clinical practice in the Oxford University Hospitals (OUH) NHS Trust and other centres around the UK. This model evaluated the costs and outcomes associated with five comparators, considered to be the most likely current and future testing strategies:

Do not share. Please treat the information presented in this supporting document as confidential

- Comparator A: current practice in the OUH NHS Trust, with genetic testing guiding treatment decisions.
- Comparator B: as per comparator A, but with Ibrutinib treatment in place for patients who relapse.
- Comparator C: current practice in hospitals which do not use genetic testing to guide treatment decisions.
- Intervention 1: genomic testing used to divide patients into likely FCR-responders (who receive Ibrutinib as refractory treatment) and non-responders (who receive Ibrutinib as first-line treatment).
- Intervention 2: as per Intervention 1, but non-responders only receive Ibrutinib as refractory treatment.

The cycle length in the model was 28 days, and the time horizon was 30 years. Patients entered the model aged 65 years. Data on the efficacy of current genetic testing practice and future genomic testing practice was provided by a retrospective sample analysis study conducted at the University of Oxford, using DNA samples from clinical trial patients. Model transition probabilities were derived from published clinical trials. Resource use in different health states was estimated via interviews with clinicians, and included clinical procedures, surgeries, hospitalisations, diagnostic tests, chemotherapy treatments, concomitant medications, complications and the time of clinical staff. Unit costs were generally extracted from published sources. Genetic and genomic test costs were estimated via a microcosting study. Ibrutinib is not licensed for use in the UK, so no price data was available. An annual cost of treatment of £30,000 was assumed (around half the current cost in the USA). Health outcomes (utilities) were extracted from published sources which used preference-based approaches within general population samples. A half-cycle correction was applied, and all costs and outcomes were discounted at 3.5%.

When the analysis was extended to a societal perspective the following cost categories were also considered: short-term absenteeism and presenteeism at work, long-term absenteeism at work due to CLL-related morbidity, productivity losses due to CLL-related mortality, informal care costs, and out of pocket costs for patients and carers.

Cohort analysis was used to conduct the economic evaluations. One-way, scenario and probabilistic sensitivity analyses (PSA) were conducted. The results of the PSA were used to estimate cost-effectiveness acceptability curves (CEACs) and calculate the expected value of perfect information (EVPI).

Cost-benefit analysis (CBA) informed by a discrete choice experiment (DCE)

A DCE was conducted to generate data on patient willingness-to-pay (WTP) for genomic testing, which then fed into the CBA. In this DCE, participants were faced with 16 choice scenarios containing two hypothetical configurations of genetic or genomic testing in CLL. By selecting one of these test configurations in each scenario, participants made trade-offs between different test characteristics (attributes). This revealed the relative strength of their preferences for different combinations of attributes and attribute levels.

The DCE was conducted amongst all CLL patients in the OUH NHS Trust (n=140) and all patient members of the UK CLL Support Association (n=982). Attributes and levels were identified using literature searches and interviews with clinicians and patients, with six attributes included in the DCE: time to receive test result, test cost, proportion of FCR non-responders identified, test reliability (how often non-responders are mis-identified), time taken to describe the test to patients, and type of clinician who explains the test result to patients. A D-efficient fractional factorial DCE design was produced using a model averaging approach, incorporating information on coefficient priors from pilot studies. Several models were considered within the data analysis, including multinomial logit, mixed logit and latent class models. The mixed logit model was found to be most appropriate in this context. The results of this model were used to calculate overall patient WTP for current genetic and future genomic testing practice.

These WTP estimates were then used in a CBA which adopted a societal perspective. As the DCE established the value that patients place on possessing genetic information, the costs included in the CBA are those which are incurred (by healthcare providers and patients) to acquire this information. Consequently, the CBA has a narrower scope than the CEA and CUA, excluding both the costs of treatment following testing and a monetary valuation of the benefits accrued following testing. Given this, a benefit transfer analysis was undertaken to establish how high WTP for treatment outcomes following testing would need to be for genomic testing to yield a net benefit.

RESULTS (a summary can be found at the end of this section)

CEA and CUA

The CEA and CUA results are presented in the first table on page three. Comparator B yields the most life years per patient while Intervention 1 yields the highest QALY gain. The most expensive strategy from both perspectives is Intervention 1. In both the CEA and CUA, the most cost-effective strategy is Comparator A (current practice in the OUH NHS Trust) when compared to the standard NICE threshold of £30,000 per LY/QALY gained. Few of the parameter variations considered in the sensitivity analysis changed the most cost-effective strategy from Comparator A to an alternative. One exception was the cost of Ibrutinib treatment: when this was reduced by a

Do not share. Please treat the information presented in this supporting document as confidential

third, Intervention 1 became the most cost-effective strategy in the CUA. The CEACs confirmed that the most cost-effective strategy at the NICE threshold was Comparator A. Assuming a 10 year time horizon and taking a healthcare provider perspective, EVPI peaks at a threshold of £35,000 in the CEA and £60,000 in the CUA (at values of £20 million and £68 million, respectively). Adopting a societal perspective yields similar EVPI results.

	Comparator	Mean LYs or QALYs per patient	Mean costs per patient (HP perspective)	Mean costs per patient (SO perspective)	ICER (HP perspective) *	ICER (SO perspective) *
CEA	C	6.37	£69,704	£73,832	-	-
	A	6.61	£71,576	£76,035	£7,903	£9,302
	Intervention 2	6.65	£91,790	£95,031	EXT.DOM	EXT.DOM
	B	7.63	£107,703	£110,820	£35,376	£34,062
	Intervention 1	7.45	£119,088	£122,116	DOM	DOM
CUA	C	5.60	£69,704	£73,832	-	-
	A	5.82	£71,576	£76,035	£8,565	£10,081
	Intervention 2	5.93	£91,790	£95,031	EXT.DOM	EXT.DOM
	B	6.44	£107,703	£110,820	EXT.DOM	EXT.DOM
	Intervention 1	6.67	£119,088	£122,116	£55,891	£54,207

* Extendedly dominated strategies (with an ICER that is higher than the next most effective strategy) were dropped in this analysis as combinations of different strategies might be feasible, given differences in current practice in the UK at present. DOM = dominated; EXT.DOM = extendedly dominated; HP = healthcare provider; ICER = incremental cost-effectiveness ratio; LYs = life years; QALYs = quality adjusted life years; SO = societal.

DCE and CBA

219 CLL patients completed the DCE (response rate 20%). Patients preferred tests that identified more FCR non-responders, were more reliable, cheaper, and returned results quickly, preferably in a 15 minute appointment with a CLL specialist. Both clinical and process-related attributes were important to patients, with ability to identify FCR non-responders and speed of test result ranked highly. Patients were willing to make notable trade-offs for the extra information provided by genomic testing, including waiting a month longer for test results. As per the table below, WTP for genomic testing was £401 higher than WTP for current genetic testing practice. This was mainly driven by the improved reliability of genomic testing. When costs and benefits were combined, genomic testing had the highest net benefit (WTP minus costs). Furthermore, the 95% confidence intervals for the net benefit difference between genetic and genomic testing did not span zero.

	Cost	WTP	Net benefit / loss (95% confidence interval)
Genetic test	£281	£254	-£27 (-£1,425 to £1,468)
Genomic test	£319	£656	£336 (-£814 to £1764)
Difference	£38	£401	£363 (£149 to £724)

The benefit transfer analysis (see table below) indicated that when the full costs of testing were considered (including treatment costs), WTP for a single life year would need to be between £14,000-£16,019 for one of the two genomic testing interventions to have a positive net benefit. A review of the literature identified estimates of WTP for a year of life in other studies evaluating cancer treatments. The highest WTP value identified was £8,728.

Comparator	Societal costs per patient *	Life years *	# tests per patient ^{a, b}	Total WTP for testing ^c	WTP gap ^d	WTP per LY required ^e
Intervention 1	£122,116	7.45	5.36	£2,784	£119,332	£16,019
Intervention 2	£95,031	6.65	3.79	£1,993	£93,039	£14,000

* Derived from Markov model; ^b WTP for one genomic test is £656; ^c Tests/patient x £708; ^d Societal costs minus total WTP for testing; ^e Amount patients must be WTP for 1 life year for overall WTP to be greater than overall societal costs.

SUMMARY

Both the CEA and CUA suggest that Comparator A is the most cost-effective strategy, regardless of the analytical perspective. The DCE results suggest that patients place a higher value on genomic testing than genetic testing in this context, and the CBA results indicate that there is a net benefit associated with genomic testing. However, the benefit transfer analysis suggests that this net benefit may disappear if the scope of the CBA is widened to consider the costs and benefits accrued following testing.

Appendix 40: Survey questions

Introductory text

This short survey aims to elicit your views on which economic evaluation results you believe best capture the true costs and benefits of genomic testing. The clinical context for this survey is the use of targeted next generation sequencing to guide chemotherapy treatment decisions for patients with chronic lymphocytic leukaemia (CLL).

All survey respondents received an email invitation to participate in this survey. As part of this invitation, you were also sent a document which summarises the study background, methods and results. Please make sure that you have read this document carefully before beginning this survey.

The survey should take 10-15 minutes to complete. Questions marked with a red asterisk (*) are compulsory.

Please treat the information presented in the supporting document as confidential. This work has not yet been submitted for publication.

This survey is part of a larger body of work investigating the methodological challenges associated with conducting health economic evaluations of genomic tests in the UK NHS. This work is being undertaken at the Health Economics Research Centre (HERC) at the University of Oxford, funded by the National Institute for Health Research. More information can be found via [this link](#).

The results of this work will not only inform my doctoral research, but will also feed into the work that we are undertaking at HERC as part of the UK 100,000 Genomes Project.

THANK YOU IN ADVANCE FOR YOUR HELP

James Buchanan (james.buchanan@dph.ox.ac.uk)
Health Economics Research Centre (HERC), University of Oxford

Questions

1. Please enter your name. This information will only be used to track who has responded to our survey. Your comments will remain anonymous. If your responses are subsequently reported in a peer-reviewed publication, we will redact any comments that could identify you as the responder.
2. How would you categorise your knowledge of cost-effectiveness analysis and cost-utility analysis?
 1. I have been the lead health economist on one or more cost-effectiveness or cost-utility analyses
 2. I have been involved in one or more cost-effectiveness or cost-utility analyses, but not as the lead health economist
 3. I have limited experience of conducting cost-effectiveness or cost-utility analyses
 4. I have no experience of conducting cost-effectiveness or cost-utility analyses
3. How would you categorise your knowledge of cost-benefit analysis?

1. I have been the lead health economist on one or more cost-benefit analyses
2. I have been involved in one or more cost-benefit analyses, but not as the lead health economist
3. I have limited experience of conducting cost-benefit analyses
4. I have no experience of conducting cost-benefit analyses

Imagine that you have been asked to consider whether the use of genomic testing to guide chemotherapy treatment decisions for CLL patients is a good use of limited resources in your healthcare system.

To help you to make this decision, you have been provided with supporting materials that summarise the methods and results of several economic evaluation approaches.

Now, please answer the following questions.

4. The results of five economic evaluation approaches have been presented. Please select the approach that you believe best capture the costs and benefits of genomic testing in CLL. More than one option can be selected, if desired.
 1. Cost-effectiveness analysis (NHS perspective)
 2. Cost-effectiveness analysis (Societal perspective)
 3. Cost-utility analysis (NHS perspective)
 4. Cost-utility analysis (Societal perspective)
 5. Cost-benefit analysis, informed by the discrete choice experiment
 6. I have not been provided with enough information to answer this question
 7. I don't believe that any of these results capture the true costs and benefits of genomic testing in CLL
5. Please explain why you selected the economic evaluation approach(es) in the previous question. If you selected option 6 or option 7, please describe what additional information you would require to be confident that the true costs and benefits of genomic testing have been captured.
6. What adoption decision would you recommend in this case?
 1. Adopt genomic testing
 2. Adopt genomic testing, but with some conditions e.g. for specific patient subgroups, only after later lines of treatment
 3. Delay a decision until additional research is undertaken
 4. Do not adopt genomic testing
7. Please explain why you would make this decision? If you selected option 2, what conditions would you consider to be appropriate in this case? If you selected option 3, what additional research would you require to be undertaken before you would be willing to make a recommendation on adoption?
8. In your opinion, would it be helpful for healthcare decision-makers at the national level (e.g. NICE, CADTH, PBAC) to consider cost-benefit analysis results alongside cost-effectiveness or cost-utility analysis results when determining whether or not to adopt genomic testing for CLL patients?

1. Yes
 2. No
 3. Unsure
9. Please explain the reasoning behind your response to the previous question. Would you change your response if the question referred to healthcare decision-makers at a more local level (e.g. clinical commissioning groups in the UK NHS)?

If your answer to Question 8 was Yes, please go to Question 10. If your answer to Question 8 was No or Unsure, please go to Question 11.

10. Are there any other circumstances in which it would be helpful for healthcare decision-makers to consider cost-benefit analysis results alongside cost-effectiveness or cost-utility analysis results when deciding whether or not an intervention should be adopted in a healthcare system?

Please go to Question 12.

11. Are there any circumstances in which it would be helpful for healthcare decision-makers to consider cost-benefit analysis results alongside cost-effectiveness or cost-utility analysis results when deciding whether or not an intervention should be adopted in a healthcare system?
12. Genomic technologies are yet to be implemented widely, in part due to the lack of evidence to support their use, particularly economic evidence. Some researchers believe that there are a number of challenges for health economists in this context, including the need to capture information on non-health outcomes (suggesting a potential role for cost-benefit analysis), the heterogeneity of cost data, and how to incorporate complex effectiveness data into economic analyses. More information about these methodological challenges can be found [here](#).

Do you think that standard health technology assessment methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? As a guide to standard health technology assessment methods, a summary of the NICE Reference Case can be found [here](#).

1. Yes [\[Go to 13\]](#)
 2. Only in specific circumstances e.g. for certain types of genetic or genomic test [\[Go to 14\]](#)
 3. No [\[Go to 14\]](#)
13. Please explain how you would like to see these methods adapted. [\[End of survey\]](#)
14. Please explain the reasoning behind your response to the previous question. [\[End of survey\]](#)

End message

You have reached the end of the survey. Thank you for taking the time to participate.

If you would like to be notified about the survey results, please email James Buchanan (james.buchanan@dph.ox.ac.uk).

Further information about this work can be found on the Health Economics Research Centre website [here](#).

Screenshots of survey questions

LimeSurvey

What is the most appropriate approach to use when conducting economic evaluations of genomic diagnostic tests?

This short survey aims to elicit your views on which economic evaluation results you believe best capture the true costs and benefits of genomic testing. The clinical context for this survey is the use of targeted next generation sequencing to guide chemotherapy treatment decisions for patients with chronic lymphocytic leukaemia (CLL).

All survey respondents received an email invitation to participate in this survey. As part of this invitation, you were also sent a document which summarises the study background, methods and results. Please make sure that you have read this document carefully before beginning this survey.

The survey should take 10-15 minutes to complete. Questions marked with a red asterisk (*) are compulsory.

Please treat the information presented in the supporting document as confidential. This work has not yet been submitted for publication.

This survey is part of a larger body of work investigating the methodological challenges associated with conducting health economic evaluations of genomic tests in the UK NHS. This work is being undertaken at the Health Economics Research Centre (HERC) at the University of Oxford, funded by the National Institute for Health Research. More information can be found via [this link](#).

The results of this work will not only inform my doctoral research, but will also feed into the work that we are undertaking at HERC as part of the 100,000 Genomes Project.

THANK YOU IN ADVANCE FOR YOUR HELP

James Buchanan (james.buchanan@dph.ox.ac.uk)

Health Economics Research Centre (HERC), University of Oxford

Next ➔

Load unfinished survey

Exit and clear survey

LimeSurvey

What is the most appropriate approach to use when conducting economic evaluations of genomic diagnostic tests?

0% 100%

Please enter your name.

This information will only be used to track who has responded to our survey.

Your comments will remain anonymous. If your responses are subsequently reported in a peer-reviewed publication, we will redact any comments that could identify you as the responder.

How would you categorise your knowledge of cost-effectiveness analysis and cost-utility analysis?

Choose one of the following answers

- ☐ 1. I have been the lead health economist on one or more cost-effectiveness or cost-utility analyses
- ☐ 2. I have been involved in one or more cost-effectiveness or cost-utility analyses, but not as the lead health economist
- ☐ 3. I have limited experience of conducting cost-effectiveness or cost-utility analyses
- ☐ 4. I have no experience of conducting cost-effectiveness or cost-utility analyses

How would you categorise your knowledge of cost-benefit analysis?

Choose one of the following answers

- ☐ 1. I have been the lead health economist on one or more cost-benefit analyses
- ☐ 2. I have been involved in one or more cost-benefit analyses, but not as the lead health economist
- ☐ 3. I have limited experience of conducting cost-benefit analyses
- ☐ 4. I have no experience of conducting cost-benefit analyses

Previous Next

Resume later

Exit and clear survey

What is the most appropriate approach to use when conducting economic evaluations of genomic diagnostic tests?0%  100%

Imagine that you have been asked to consider whether the use of genomic testing to guide chemotherapy treatment decisions for CLL patients is a good use of limited resources in your healthcare system. To help you to make this decision, you have been provided with supporting materials that summarise the methods and results of several economic evaluation approaches.

Please answer the following questions.

The results of five different economic evaluation approaches have been presented.

Please select the approach that you believe best capture the costs and benefits of genomic testing in CLL.

More than one option can be selected, if desired.

Check any that apply

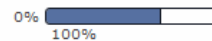
- ☐ 1. Cost-effectiveness analysis (NHS perspective)
- ☐ 2. Cost-effectiveness analysis (Societal perspective)
- ☐ 3. Cost-utility analysis (NHS perspective)
- ☐ 4. Cost-utility analysis (Societal perspective)
- ☐ 5. Cost-benefit analysis, informed by the discrete choice experiment
- ☐ 6. I have not been provided with enough information to answer this question
- ☐ 7. I don't believe that any of these results capture the true costs and benefits of genomic testing in CLL

Please explain why you selected the economic evaluation approach(es) in the previous question.

If you selected option 6 or option 7, please describe what additional information you would require to be confident that the true costs and benefits of genomic testing have been captured.

[Previous](#) [Next](#)[Resume later](#)[Exit and clear survey](#)

What is the most appropriate approach to use when conducting economic evaluations of genomic diagnostic tests?



* **What adoption decision would you recommend in this case?**

Choose one of the following answers

- ☐ 1. Adopt genomic testing
- ☐ 2. Adopt genomic testing, but with some conditions e.g. for specific patient subgroups, only after later lines of treatment
- ☐ 3. Delay a decision until additional research is undertaken
- ☐ 4. Do not adopt genomic testing

* **Please explain why you would make this decision?**

If you selected option 2, what conditions would you consider to be appropriate in this case?

If you selected option 3, what additional research would you require to be undertaken before you would be willing to make a recommendation on adoption?

* **In your opinion, would it be helpful for healthcare decision-makers at the national level (e.g. NICE, CADTH, PBAC) to consider cost-benefit analysis results alongside cost-effectiveness or cost-utility analysis results when determining whether or not to adopt genomic testing for CLL patients?**

Choose one of the following answers

- ☐ 1. Yes
- ☐ 2. No
- ☐ 3. Unsure

*
In your opinion, would it be helpful for healthcare decision-makers at the national level (e.g. NICE, CADTH, PBAC) to consider cost-benefit analysis results alongside cost-effectiveness or cost-utility analysis results when determining whether or not to adopt genomic testing for CLL patients?

Choose one of the following answers

- ☐ 1. Yes
☐ 2. No
☐ 3. Unsure

*
Please explain the reasoning behind your response to the previous question. Would you change your response if the question referred to healthcare decision-makers at a more local level (e.g. clinical commissioning groups in the UK NHS)?

*
Genomic technologies are yet to be implemented widely, in part due to the lack of evidence to support their use, particularly economic evidence. Some researchers believe that there are a number of challenges for health economists in this context, including the need to capture information on non-health outcomes (suggesting a potential role for cost-benefit analysis), the heterogeneity of cost data, and how to incorporate complex effectiveness data into economic analyses. More information about these methodological challenges can be found [here](#).

Do you think that standard health technology assessment methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? As a guide to standard health technology assessment methods, a summary of the NICE Reference Case can be found [here](#).

Choose one of the following answers

- ☐ 1. Yes
☐ 2. Only in specific circumstances e.g. for certain types of genetic or genomic test
☐ 3. No

[Previous](#) [Submit](#)

[Resume later](#)

[Exit and clear survey](#)

LimeSurvey

What is the most appropriate approach to use when conducting economic evaluations of genomic diagnostic tests?

You have reached the end of the survey. Thank you for taking the time to participate.

If you would like to be notified about the survey results, please email James Buchanan (james.buchanan@dph.ox.ac.uk).

Further information about this work can be found on the Health Economics Research Centre website [here](#).

Appendix 41: Survey of healthcare decision-makers – responses

Table 1: Responses to questions regarding the appropriate economic evaluation approach and the recommended adoption decision

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt	
s1						✓		The correct method depends upon the budget that will provide funding for the test and associated treatment. An NHS perspective is consistent with a wide range of budgets, some of which it would be appropriate to use CEA and others it would be appropriate to use CUA. Once might consider CBA as an appropriate method if the funds were coming from the broader public purse budget, rather than the NHS budget. The correct method depends upon inter alia (a) the objective of the decision-maker the analysis is undertaken to inform, and (b) the budget that the decision-maker is allocating. This information is not provided.			✓		Again, insufficient information has been provided to answer this question.
s2						✓		The "best" approach to use is that which is most legitimate. If the decision-maker is delegated an objective by a legitimate higher authority (e.g. parliament), then this objective ought to state the required perspective and the desired outcome measure. Whether I (as an analyst) prefer to use QALYs or another measure of effectiveness, or prefer an NHS or societal perspective, is irrelevant. Without being informed of the objective in question I cannot answer the question.			✓		I would need to consider the uncertainty around the provided results before making a decision. Knowing that the CEAC is >50% for Comparator A at a £30,000/QALY threshold is insufficient.
s3		✓	✓					I require option number 3 in order to prioritise investments from a fixed health budget. Assuming it is not cost effective enough to be funded on this basis, I would require option 2 in order to make an effective budget bid for additional resources from Cabinet.			✓		Question is poorly worded, I have not seen the actual results of the analysis, and I am unaware of my budget constraint, so a recommendation at this time is nonsensical. I could adopt any of the recommendations given and justify it here. So, on that basis, I have chosen my answer as the results I saw discounted at the wrong rate. If the result had been done exactly occurring to methods and was cost saving, I would adopt.

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt	
s4				✓				I think it is difficult to separate testing from the consequences of testing in terms of treatment, which is the basis for rejecting the CBA approach. CUA over CEA as prefer quality to be included however imperfectly.		✓			My condition would be to gather all the information prospectively and centrally, to ensure availability for future research.
s5			✓					If using NHS resources the opportunity costs fall on health. Need a generic measure to compare health gained and health lost. If wider perspective then need to (a) specify social welfare (we don't know) and (b) assess the other non-health benefits foregone (we generally don't know).	✓				Comparator A: Don't have enough information to decide if more research needed, what type, what irrecoverable costs and therefore whether would need to delay adoption until research conducted.
s6			✓					The spend will fall to NHS budgets and so the opportunity cost is other NHS spend.	✓				ICER versus current practice is favourable.
s7						✓	✓	I have little confidence in the results. First, the WTP to avoid death is implausibly low. UK estimates of WTP to extend life are generally in excess of £30,000/LY (Mason <i>et al.</i> 2009; Donaldson <i>et al.</i> 2011). Second, the analysis should have compared an intervention with NGS and FCR drugs with the comparator A, and not confound genomic testing with a new drug. It would seem likely that a genomic test that costs little more than the genetic test it replaces but is twice as accurate would be much more cost-effective if not cost-saving. Why was that strategy not modelled? When that is combined with the DCE results, genomic testing might dominate all alternatives.			✓		Because I don't consider the analysis that was presented to be sufficiently informative.
s8							✓	I think that budget impact analysis will often be an important part of an overall assessment.			✓		It would be helpful to better understand how clinically important it is in this case for patient preferences (utilities or WTP) be included. Many testing decisions are more focused on provider decision-making and outcomes that may not be captured by utilities or DCE.

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt	
s9				✓		✓		I would require a transparent process for agreeing for a methodology to determine "value for money" in my health care system which has not been achieved yet. In case of doubt, I think CUA from societal perspective is most likely to be the agreed methodology.		✓			Assuming that there is good evidence of effectiveness, I'd recommend adoption; otherwise some managed entry agreement.
s10	✓	✓						I find measures of effectiveness more robust and more easy to defend than utility or monetary measures.		✓			Evidence presented.
s11	✓	✓	✓	✓				In combination these analyses provide multiple estimates of benefits (LY and QALY) and multiple cost perspectives, allowing a reasonable characterisation of the effect on the ICER of inclusion/exclusion of different data.				✓	Comparators A and B (with genetic, not genomic, testing) seem to offer better value for money relative to current practice.
s12					✓			The question was to identify that which best captures the costs and benefits of genomic testing, rather than which one should be used in decision-making. In theory the DCE should elicit a WTP where rational respondents take all opportunity costs into account, therefore this captures the most costs and benefits. This does not mean it's the most useful for informing adoption decisions in the NHS though.				✓	The treatment itself is not cost-effective (I'm assuming so anyway as it is covered by the CDF), so therefore it is unlikely for a diagnostic to provide access to that treatment to be cost-effective. However, where the diagnostic can ensure treatment is given only to those who are more likely to benefit, the overall strategy may be cost-effective. This does not appear to be the case here though.
s13	✓		✓					Decision-making for the NHS has been standardised by NICE as a CUA. While not necessarily ideal for genomic testing in CLL it does provide a level of comparability for appropriate allocation of resources across a range of areas.				✓	No information is provided upon which to select one of the answers? In the absence of any evidence suggesting its clinical benefit (let alone its cost-effectiveness) I would not recommend adopting it.
s14						✓		The answer to this question is dependent on the research question that is being answered. The type of analysis that is chosen needs to reflect the research question that is informed by whom is making the decision and the reason for conducting the analysis.		✓			On the basis of the information presented genomic testing cannot be introduced on a widespread basis after this study. Some key aspects were not presented such as how clinicians should/do use the information to guide treatment selection.

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt	
s15			✓	✓				CUA is preferred to CEA because the QALY is a broader measure of outcome than the LY. The DCE and hence the CBA provide a limited view of the benefits of testing.			✓		Based solely on the ICER. Decisions can always be delayed, so if new data are collected (or the prices of the tests or drugs come down) which lowers the ICER, testing could then be adopted.
s16		✓		✓				Decisions around scarce resources should include all costs and effects/utilities (both direct and indirect).		✓			Based on evidence provided. Might be even more cost-effective if stratified by specific groups etc.
s17	✓	✓	✓	✓	✓			Each of the analyses provide slightly different information. In general CEA and CUA do not capture important preference information as does DCE, however the latter is not easily interpreted by policy decision-makers.			✓		In general the information favours genetic testing, however the chance that there may not be a net benefit if follow-up costs are considered suggests a potentially unacceptable risk.
s18					✓			On the basis of a very limited understanding of the details of cost-benefit analysis and DCE i.e. only what I have read in the handout, it seems to me that an attempt to get an understanding of the kinds of choices people would make in practice and why - and what matters to them - is not only an important way of potentially identifying unexpected factors of relevance - and hence a potential discipline on the assumptions made in the cost-benefit analysis but also and relatedly a useful way of allowing for the taking into account of the fact that people are not (only) rational choosers.		✓			Adopt genomic testing but in parallel with health economic and social scientific research/evaluation so that learning can go on in parallel and be reviewed at a later stage. Doing research below implementation will inevitably continue to be somewhat speculative.

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt	
s19	✓		✓			✓		For option 6, part of my issue is that my background is primarily in psychology, so the limitations of DCEs (in terms of realism, cognitive burden, etc.) means it's hard for me to assess the value. I'd prefer more information on how/why attributes were selected, what information was gleaned from qualitative work vs the literature, etc. I'd also like more information about the assumptions guiding the economic evaluations. My reasons for preferring the NHS perspective are due to this being the most "concrete"; though in general I think societal views are more useful.	✓				From my perspective, the current NHS guidelines are valid.
s20						✓		I'd need to know the decision environment, are we to take into account standard HTA? If not then a CBA, but one where the NHS and societal perspective can be presented separately.			✓		Given some of the net benefits appear to disappear, I think it would be important to understand why.
s21				✓				To me, most transparent and comprehensive in terms of costs and utility of testing + treatment consequences.			✓		Delay until ibrutinib treatment becomes cheaper and/or decision uncertainty is reduced. The EVPI is still substantial and it would be helpful to have ENBS estimates to decide on further research.

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e	
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt		
s22							✓	I am not entirely clear of the composition of the genomic panel used (i.e. is it specific just to CLL mutations, does it include germline and somatic mutations, etc.). I believe that this is an important consideration as the benefits derived from such testing are largely dependent on the composition of the test. For example, if the genomic test included information that would be useful for the patient outside of CLL I don't think the analyses have captured these benefits. The value in genomic testing is not for one off treatment decisions but for a number of health related decisions that will happen over a patient's life time with only a single test required to derive this information. If the genomic panel is just a targeted CLL panel composed only of genes commonly mutated in the disease and no other genomic information is collected I think the CBA (net benefit from difference in cost of genomic test versus WTP) is the best measure of the value of just the test. I say this because the results of the CUA and CEA I would assume are largely driven by the costs of the expensive therapy. If the test is used to select a cost-ineffective therapy then more than likely the test will not be cost-effective. I do agree however that genomic testing needs to be assessed with treatment but I think standard CEA and CUA methods bias results against genomic testing especially in cancer where the associated treatment is usually quite expensive.		✓				It is difficult to be very prescriptive in what conditions to apply as I am not sure of the quality of the clinical evidence and how different these populations are to the model population but I think you would want to know is this extra testing resulting in better overall survival and are clinicians actually using this information in their decision-making (clinical utility). Access to testing should not be delayed entirely but evidence should be collected as described in the question.
s23				✓				This is the most comprehensive and comparable evaluation.				✓	Taking the NHS view, genomic testing is not considered to be cost-effective. Further, the EVPI suggests that a future trial is unlikely to be worth investment (EVSI for a trial would likely be less than the EVPI).	

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt	
s24				✓				CUA, because it is important to be able to determine the opportunity cost across the entire health domain rather than just within the context of a single condition. Also since CEAs often only measure interim outcomes and do not necessarily measure health gain (rather health gain by proxy) they are not sufficient for high level decision making (e.g. public funding). Societal perspective, where possible this approach should be included since it provides a more comprehensive picture of the costs and benefits. This is important when evaluating decisions regarding public funding since this funding is taken from a poll that is not limited to health. However I feel that the analysis should also report the NHS perspective - particularly when determining the financial impact.				✓	Given current evidence as presented the intervention is not cost effective compared to current practice.
s25						✓		The key driver in my view is the efficiency question to be addressed. If it's technical it is CEA, if allocative then CBA. It has nothing to do with how outcomes are measured and valued. The idea of societal CEA, for example, is a nonsense.			✓		I would tick any of the answers really. The way most health economic evaluations are done, however, led me to tick the third option because I'd want to know the opportunity cost.
s26				✓	✓			CUA and CBA are approaches that can incorporate health and other consequences.			✓		The obvious answer is, I do not know, as the details of the VOI analysis are not provided. In general, more knowledge of identifying non responders on (mostly very expensive) targeted therapies is usually valuable. Often even for other populations than the one that is researched specifically.

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt	
s27			✓					Prefer CUA to CEA because survival doesn't necessarily capture the full quality of life benefits that may accrue to patients (but recognising many issues associate with applying utility weights); Prefer a health system perspective to a societal one because of the risks of the uncertainties that can be introduced by the latter; A genomic or genetic test is only 'valuable' if it changes subsequent patient management, and those changes are associated with improved health outcomes (LYs, QALYs gained, or toxicity avoided). The CBA is therefore difficult to interpret because the patient benefit is assumed (via number of likely non-responders identified), but is not explicit.			✓		Whilst the ICER associated with Intervention 1 is above the cost-effectiveness threshold, the QALY gains with Intervention 1 are substantial. I would want to explore the effect of employing the genomic test in a selected patient population, perhaps defined by family history, results from previous genetic tests, other biomarkers etc.
s28			✓	✓				I've selected CUA (NHS) as I believe that this is what decision-makers need to draw comparisons of cost-effectiveness across different treatments/clinical conditions. However, I've also ticked CUA societal as I think it is important that both are presented for decision-makers to see the magnitude of costs (and benefits) falling outside the health care sector (although in the scenario presented only costs were measured from societal, not benefits). I didn't tick CBA as I think it's important to include costs and benefits after testing so this felt like a major limitation.				✓	Regardless of the analytical perspective (for CEA and CUA), comparator A was the most cost-effective strategy and the ICER was well under the £20k threshold.

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt	
s29	✓	✓	✓	✓	✓			There are concerns to be highlighted regarding the low response rate of the DCE and the potential for hypothetical bias. Also, as the authors know, the cost of ibrutinib is quite high. A priori, without a change in list price, I would be surprised if there is a situation where ibrutinib in CLL would be cost-effective. As such, public engagement surrounding its coverage may be warranted; the DCE is a first step for consulting the public. I did think it an interesting approach regarding estimating the 'value of knowing' and the cost-benefit of providing knowledge followed by a threshold type analysis looking at the needed WTP for a single life year for an overall positive net-benefit. In terms of the highest WTP value identified from the literature review, this looks a wee bit low to me, but then again I didn't conduct the literature review so difficult to comment. In general, benefit transfer analysis is an interesting angle in this arena. In all, I would guess that a pragmatic approach is warranted; that is, I would take all the analyses and put them in front of decision-makers. I'd then probably suggest an MCDA-type approach to decision-making. Further, as noted above, public engagement may be an avenue to explore. If I were forced to 'pick one or two' analyses, I would go with the CUA NHS perspective and the CBA.			✓		From what I've been told, the clinical research surrounding ibrutinib in CLL needs more time to convincingly demonstrate long term benefit in terms of life years and quality of life. Further research including a cost-effectiveness component (or alongside cost benefit analysis!) is an important area of future research given the high cost of the drug.
s30				✓				Comparable and comprehensive.				✓	Current practice is the most cost-effective.
s31		✓		✓				With my limited knowledge I feel that either analysis needs to include wider societal perspective - the NHS is increasingly adopting a patient-centric / personalised medicine approach and costs must reflect this.	✓				It feels like a more rational / ethical approach given the rate of non-response and the side effects associated with chemotherapy. I wasn't sure if Option 3 allows the improved therapy to be offered as part of a clinical trial (or NICE conditional approval) in which case this would be my preference.

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt	
s32	✓		✓					These are the methods I am most familiar with and there was insufficient additional information for the other approaches to indicate to me their greater advantage in answering the decision question.			✓		I would like to know the results of how genomic testing in current practice and guiding treatment compared to comparator A? This would answer the question of the performance of the new test versus the current test without the addition of the new treatment. Could this provide useful cost-effective evidence to support replacement?
s33						✓		I chose number 6 because I think that having access to the 16 choice scenarios would have allowed me to better understand the CBA (vs CEA/CUA presented here) and therefore make a choice between whether I think that the CBA captured the true costs and benefits of genomic testing. I also would have liked a figure showing the three comparators and two interventions as I had to read it a couple of times to understand how they all relate to each other. I should just point out that I am leaning towards the CBA though.	✓				The decision rule for CBA is straightforward - there is a statistically significant positive net benefit. CBA also allows the potential to include non-health benefits/dis-benefits which are relevant for genomics.
s34							✓	The cost, benefit and utility of genomic vs. genetic testing remain unclear as all comparisons involving genomic testing also include the use of ibrutinib, which is not currently licensed in the UK. How were benefits of avoiding side effects and cost savings of non-treatment (FCR) for non-responders identified through genomic testing accounted for? Why was the DCE conducted amongst all patients rather than those eligible for FCR (and hence genetic and genomic testing)?			✓		As per my previous response, it would in my view be useful to differentiate the decision about genetic vs. genomic vs. no testing from treatment decisions for any identified FCR non-responders.

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt	
s35						✓	✓	I would start with something more descriptive like cost-consequence and work from there to see if all the important benefit pieces could be adequately captured in CUA or CBA.				✓	I'm much more influenced by the CUA results than the CBA – patient values, only 20% response, not convinced that DCE/WTP works well in this context, etc. If there are missing benefit pieces in the CUA, which I'm sure there are, I'd prefer to see them clearly articulated and presented in the form of a CCA first, and then think about value tasks subsequently.
s36	✓							The linkage of genomic testing to treatment selection and concomitant health outcomes seems most relevant, although is not always possible to obtain this data e.g. needs linkage of the diagnostic with clinical trial.	✓				It is clear that genomic testing can influence treatment selection, and I have concerns that the benefits can be lost if the diagnostic test is delayed until after some treatments fail.
s37	✓			✓				1 is most relevant for healthcare system, 4 captures all effects.			✓		Need information on how ibrutinib patients in both 1L and 2L respond based on genetic markers used to identify FCR responders.

Resp	Please select the approach(es) that you believe best capture the costs and benefits of genomic testing in CLL							Please explain why you selected this (these) approach(es) ^c	What adoption decision would you recommend? ^d				Please explain why you would make this adoption decision ^e
	CEA (NHS)	CEA (Soc)	CUA (NHS)	CUA (Soc)	CBA	Lack of info ^a	None ^b		Adopt	Adopt with conditions	Delay decision	Do not adopt	
s38			✓					The CUA in this study takes into account a combination of LYs gained and quality of life, whereas the CEA only takes into account LYs gained. I would say the CUA approach is therefore more complete. For the CBA I would like more information on the experiment and its results. Additionally, I find it questionable whether the DCE should be done by patients or by the general population.			✓		NGS is a field that is rapidly evolving. Prices are known to have decreased a lot, and they might continue to fall even further. Additionally, the technology is still advancing considerably. Therefore NGS might be expected to result in a higher diagnostic yield at lower costs in the near future. It might be informative to model the expected near-future cost decreases and technological advancements (e.g. in terms of diagnostic yield – where WES is known to be insensitive to some types of mutation, such as CNV's, WGS will already pick these up).

CBA = cost-benefit analysis; CCA = cost-consequence analysis; CEA = cost-effectiveness analysis; CEAC = cost-effectiveness acceptability frontier; CLL = chronic lymphocytic leukaemia; CUA = cost-utility analysis, DCE = discrete choice experiment; ENBS = expected net benefit of sampling; EVPI = expected value of perfect information; EVSI = expected value of sample information; FCR = fludarabine, cyclophosphamide, rituximab; HTA = health technology assessment; ICER = incremental cost-effectiveness ratio; LY = life year; MCDA = multi-criteria decision analysis; NGS = next generation sequencing; NHS = national health service perspective; NICE = National Institute for Health and Care Excellence; QALY = quality adjusted life year; Resp = respondent; Soc = societal perspective; VOI = value of information; WES = whole exome sequencing; WGS = whole genome sequencing; WTP = willingness-to-pay. ^a The exact wording for this answer was: "I have not been provided with enough information to answer this question". ^b The exact wording for this answer was: "I don't believe that any of these results capture the true costs and benefits of genomic testing in CLL". ^c The exact wording for this question was: "Please explain why you selected the economic evaluation approach(es) in the previous question. If you selected option 6 or option 7 please describe what additional information you would require to be confident that the true costs and benefits of genomic testing have been captured". ^d The exact wording for the four options was: 1. Adopt genomic testing; 2. Adopt genomic testing, but with some conditions e.g. for specific patient subgroups, only after later lines of treatment; 3. Delay a decision until additional research is undertaken, and 4. Do not adopt genomic testing. ^e The exact wording for this question was: "Please explain why you would make this decision. If you selected option 2, what conditions would you consider to be appropriate in this case. If you selected option 3, what additional research would you require to be undertaken before you would be willing to make a recommendation on adoption?".

Table 2: Responses to questions regarding the appropriate methodologies to use in genomics

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s1	Y	I believe that some form of information on the opportunity cost of new technologies should be considered by decision-makers allocating limited budgets. Which type of information is most useful depends upon the objective of the decision-maker and the budget they are allocating. I believe this to be true irrespective of the level within the system the decision-maker operates, subject to the constraint that the resources consumed in providing the information are proportionate to the scale of resources being allocated.	If decision-makers understand their objective and the budget they are allocating, they should be able to choose the most appropriate form of information. To present all forms of economic evaluation may well introduce confusion to no clear advantage.	Y	But not for the reasons advanced in the paper cited. The value of technology that calls upon the health care budget is determined by the objective given to the health care system alongside the budget. A consistent value framework is required in order for reimbursement decisions to be internally coherent with regard to efficiency and horizontal and vertical equity concerns. The measure of value does not flow from the nature of the technology being evaluated, it flows from the objective of the health system. There are however, questions about how to optimise genomic technologies to meet value based reimbursement decision rules that are not well dealt with in the current literature.

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s2	U	It is not apparent from the information provided what the costs/benefits would be of considering CBA instead of, or in addition to, CEA or CUA.	See previous answer. One concern with using CBA as presented would be the WTP value used. The "review of the literature" may have included estimates of individual's WTP out of their own pocket to improve their own health (or WTP to avoid risk to themselves), which is not the appropriate measure of WTP for a public health system with a constrained budget, in which the opportunity cost of funding new technologies falls on unidentified member of society, and where health care is publicly financed.	N	The methods should no doubt be adapted (they are not perfect) but not specifically for genetic or genomic tests. Also there is no consensus on the "need" to capture non-health outcomes. In most public health systems, the health budget has been allocated by parliament with the (primary) purpose of improving population health. Considering non-health outcomes raises the potential that resources may be allocated in such a way that "worsens" population health but improves some individuals' quality of life in other ways. What is the legitimacy of doing this?
s3	U	It's only helpful to add additional analyses if your primary analysis does not recommend a very specific course of action. 95/100 a CUA if done well is sufficient.	When bidding to cabinet for additional resources. They will need to be able to summarise the gains to society beyond the health care gains seen, and to know the relative cost effectiveness of the interventions they have considered compared to other possible investments.	Y	In my practice these methods are already used. Key areas to focus on are: target population (can you target to a group likely to benefit more); false positives and negatives (and consequential costs and benefits); date of diagnosis (early treatment gets better results), and transition probabilities, costs, benefits.
s4	U	I need more information on how this would work in practice, and on how CBA would 'score' accepted approaches to treatment/testing. No change with respect to the local level.	Given the caveat in my previous response, yes.	Y	Access to treatments is increasingly predicated on genetic/genomic (or other) methods of categorisation. Costs of 'finding' patients must be included in assessments. The failure to do so in previous HTAs is perverse.

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s5	U	Would only be useful if one can consider the other non QALY aspects additional NHS costs are likely to displace. If not then make decisions that not only reduce health (QALYs) but also other aspects of value too.	See previous response. I don't think the real distinctions are about CEA, CBA and so called CUA or about health technologies vs genomic. It's just about what aspects of outcome are valuable, whether you can account for the opportunity costs, and what is the value of turning variability into heterogeneity. These are relevant to all health care choices, not just genomic, so should not use different methods just for them.	N	See previous answer: if there are good reasons to change methods then relevant to all not just genomic.
s6	N	As mentioned before the costs will fall to NHS budgets so that is where the opportunity cost lies.	Strictly speaking, no. Less strictly, other economic methods may be helpful to guide committee judgements.	N	Because the necessary other factors can be considered in committee.
s7	Y	National level decision makers implicitly consider the national perspective. Local decision makers are focused on budgetary impact and would likely reason differently.	Unclear.	Y	That remains to be determined.
s8	U	See previous response.	When patient preferences and net benefits are particularly important.	Y	Consideration of unique challenges such as multiple results.
s9	N	No because I believe the methods for determining WTP are not sufficiently reliable. The average of multiple stakeholders clicking quickly through a survey, potentially without understanding the issues at stake, is unlikely to improve the decision.	If the methods were better and there was a political agreement that CBA should be used.	Y	There should be the same levelling rule for every technology.

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, <u>Unsure</u>) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, <u>Only in specific circumstances</u>)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s10	U	My caution relates to how ability to pay was managed in any stated or revealed preference valuation of benefits.	Yes, potentially, if the perspective adopted was that of the health service and the valuation was based on available budgets over which they had discretion.	O	I think techniques should continue to evolve as they encounter particular challenges rather than atrophying because of a failure to adapt.
s11	U	I am simply unsure about putting a monetary value on some aspects. I think it is more helpful to count these non-health benefits/harms rather than value them e.g. in a form of cost-consequence analysis.	Maybe, but more helpful to see disaggregated 'other outcomes' rather than CBA.	O	Evaluation of any kind of diagnostic test is complex. It needs to consider not only the costs and value of the test itself (often difficult to measure in QALYs) but also the downstream costs and consequences of the alternative test/treat pathways (including consideration of TP/TN/FP/FN results). The NICE base case specifies QALYs, but I'm not sure QALYs are appropriate and/or adequate in all circumstances.

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s12	U	If the objective of a health system is to maximise length and quality of life (a big 'if' I admit), then the opportunity cost of providing genomic testing is still in terms of foregone quality adjusted survival elsewhere: the additional benefits of genomic testing (information about risk of future disease, and decisions over reproduction etc.) still have to be weighed against the foregone benefits of not providing some other life-saving or life-enhancing treatment. What is the value of the knowledge of a genetic defect in terms of patients foregoing hip replacements, lives lost in some other disease etc.?	Only if the remit of a health care system is to maximise welfare as a whole (however defined).	N	Perhaps 'no' is too strong a response, but it is very difficult for decision-makers to tell the difference between crocodile tears from industry who just want the goalposts shifted in order to maximise their returns, and the genuine limitations of the extra-welfarist approach to economic evaluation of health care interventions, namely its inability to capture everything that is important to patients. I would tend to prefer keeping the same analytic approach we have at present (speaking from an English perspective), but with supplemental analyses which may include CBA (as well as other issues outside the NICE reference case such as equity considerations, scenario analyses incorporating impact on carers etc.). This provides the information the appraisal committee needs to weigh up whether the additional elements are genuine, and thus whether the benefits of the genomic tests will truly be valued more highly by society than the loss from displaced healthcare.
s13	N	Standard metric for decision making by NICE is cost per QALY gained. Local decision-makers are typically less sophisticated than the NICE Appraisal Committee.	No - providing we are confident that all of the benefit (or harms) are adequately captured by the utility estimates.	N	Genomic technologies should not have a different/preferential treatment as the basis for decisions which would potentially reallocate resources from existing healthcare funding.

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s14	N	This would deflect the responsibility too much towards the decision-maker in terms of being able to assess whether it is CBA or CEA that is the appropriate evaluative framework to use. This would introduce the problem not currently being adequately addressed in the NICE public health programmes system that advocates using CCA and CBA and CEA - which result do you 'believe'? There is no guidance on how to rank or rate between answers produced from different evaluations and so the decision-maker is at best left to guess which they think is the right answer. No, my answer would definitely not change.	In my opinion, not at the moment - until there is some empirical work that informs a decision-maker how they SHOULD use both sets of evidence.	O	First, we need criteria to decide which technologies (and it won't just be genomic tests) cannot be appraised using current HTA methods. NICE has already done this for specialised (high cost) technologies, so there is a precedent in place. Then we need to consider how to evaluate non health outcomes but this does not suggest we SHOULD use CBA. There are other ways of capturing non health outcomes and it is first necessary to decide what definition and type of non-health outcomes are important.
s15	N	Decision-making is improving, especially with recent work around opportunity costs. Decision-makers incorporate other factors implicitly. CBA would over-complicate the process, especially given wide range of methods and design choices.	Possibly when considering alternative mode of delivery interventions, but I am generally sceptical of such studies due to risk of bias.	Y	With respect to collecting and analysing effectiveness data.
s16	U	-	-	Y	-

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s17	U	In general CBA provides more information and has additional usefulness when comparing investment intersectorally. However, the difficulties are in the interpretation. This difficulty would increase at the local level.	As mentioned in my previous response, for investment choices across sectors e.g. interventions for illicit drug use. In genetic testing, education is a large component which is a non health outcome.	Y	My experience is around undertaking cost-effectiveness of genetic tests and policy decisions concerning genetic testing. I am aware of the shortcomings of current methods and believe there is a need for better methodology such as DCEs, but have not explored the potential implications of alternate economic evaluation approaches. I look forward to outcomes of your research.
s18	Y	Again, based on limited understanding, my view is that including some evidence about real decision-makers, even if this uses a concept of WTP which is not necessarily fully relevant to NHS, is important. This may not always be useful at the local level but it does seem likely that patients will have different values in different contexts and that this might at least sometimes be relevant.	I'm not sure I am qualified to answer this one.	N	My assumption here is that CBA and DCE are not 'standard' HTA assessment methods. The reasons are those in my previous response. In addition to this I agree with the claim that non-health outcomes are important. These matter to people and are key capabilities that add to the value of life. I am wary of a societal approach which only sees this from the view of wider society e.g. the value of people in society but recognise the importance of the idea that it might be valuable to patients to be able to participate e.g. work.

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s19	U	I think this depends entirely on the ability of decision-makers to interpret the data - not British, so not fully aware of the systems in place.	Again, I am unsure as to what the NHS covers, and which costs individuals bear - but if you are calculating WTP from a DCE, the situations where I think this is most helpful are those in which individuals are dealing with semi-elective treatments (for example, general genetic counselling).	O	I think this is really dependent on the role the genetic test is playing. If it's effectively being used as a screening tool, it should be treated the way other diagnostics are. However, in circumstances where the information can be used beyond the clinical context it's collected in, I think there's a need to explore how and why genetic information may provide personal utility not captured in a standard economic evaluation.
s20	U	Essentially NICE already does a CBA, a QALY has a monetary value of 20-30k. Presenting a different money metric to them shouldn't be difficult for them to understand, but I imagine that they would question if it is measuring health, like a QALY supposedly does, or something broader.	I think this sort of evidence (if produced robustly) would be valuable to decision-makers, but they just need to understand it.	Y	I thought NICE had reviewed its reference case for genomics? Didn't Katherine Payne look at this (as well as yourselves)? Maybe they did review it, but didn't change it. It is definitely a different intervention from pharmaceuticals so needs a different reference case. The process here in Australia for co-dependents which includes genomics is different, but still flawed as the genomic testing approach drives the modelling approach, but is much less costly than the accompanying drug.

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s21	U	CBA based on DCE does give some additional insights in how patients value other aspects of testing that are not typically valued in the CEA/CUA and for other effectiveness units than QALYs or LY thresholds are unknown yet relevant. My hesitation is that results from CBAs often vary widely, depending on other factors than true patient or preference heterogeneity. Also WTP estimates often seem unreasonably (relative to health care insurance premiums or out of pocket costs that individuals or households can actually bear) high which makes me nervous about their actionability. My answer would not change if the question referred to local decision-makers. I would suggest local decision-makers to consider a budget impact analysis in addition to a CEA/CUA and/or CBA.	Yes, when non-health factors are (expected to be) of significant importance and when WTP for the effectiveness unit of choice is unknown.	Y	Need to consider how to efficiently deal with a large number of comparators to model the potential clinical applications and to determine their value. Assessing outcomes beyond health as well as specific issues, such as different payer perspectives and patient preferences, is crucial to fully capture the potential health economic impact of diagnostic biomarkers and to inform value based reimbursement. Also need to consider how much and what type of evidence is realistically needed in order to make decisions when they matter (and not years afterwards).

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s22	Y	I answered yes because as I mentioned before I think especially in cancer the value of testing by itself can be masked by the cost of the therapies it suggests. It seems the CBA result provides some insight into the value of the test alone but then decision-makers can also see the cost-effectiveness of the test and treatment together in the CEA and CUA.	If a decision-maker wanted to incorporate non-health outcomes into their decision-making process.	N	The challenges in evaluating genomic tests are not necessarily specific to genomics. The economic assessment of a number of diagnostic test and treat strategies pose similar methodological challenges. I do believe that as we start to use more comprehensive genomic tests (WGS/WES) that models will become quite complex and large but still largely believe that current methods could still be appropriately applied but the feasibility and the rapid pace at which our knowledge of genomics is growing calls into question the practicality of using current methods and potentially suggests that unique approaches might be required.
s23	N	At a practical level, it would be difficult to generate these data in addition to CUAs and also difficult for decision-makers to add this to list of things to consider. These also include benefits to patients that may be outside of the perspective of the healthcare decision-making entity.	Perhaps guidelines groups, which are beginning to consider CEA/CUA data might be interested in CBA.	N	I think the standard approaches are sufficient. It is largely an issue of data availability and not methodology.

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s24	N	I am not a fan of CBA. I feel that the methods employed to translate health outcomes into cost are not robust enough and have limited generalizability. In my experience as a national decision-maker cost utility analysis together with a financial impact analysis provides sufficient information.	Not in my opinion.	Y	I think that one of the issues with evaluating genomic tests is the lack of a gold standard - this makes measures of outcome difficult. While this is true for several diagnostic tests generally they are a replacement for a more invasive option which while not perfect is the best available. Genomic testing is unique in some cases since the only way of determining diagnostic accuracy may be to follow the patient - and then not all genetic mutations are expressed. For genomic testing that is related to suitability for a drug, there may be no concrete way of determining diagnostic accuracy since even if the drug is given to someone who should or should not receive it, the outcome would not be 100% in either direction, so this would not be a good marker of the diagnostic accuracy of the test. Due to this, measuring the incremental effectiveness of some genomic tests is difficult.
s25	U	Incremental CEA implies more resources may be allocated to the area of care/provision from which the intervention comes. Therefore, they are doing forms of CBA already, but just not fully societal.	All. If we start societal, which is what I'd always advise students to do, then CBA is the default.	N	Yes in one way. That is, it's all rubbish now and so the whole lot needs to change. If we did that, your question becomes irrelevant, which in my mind it is. Hence I ticked No.

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s26	Y	The CBA has an entirely different perspective than the CUA/CEA. The CBA focuses on the test, the CUA/CEA on the long term health outcomes. In this way, I think the CBA has very little value for decision-making. I think, for decision-making, a CBA that would combine the two perspectives: the patient value perception of the (process aspects of the) test as well as the long term consequences (being health or other). A challenge is how to avoid a focusing illusion in patients completing a DCE (or any other WTP approach) on process aspects that are relevant for a short period, and outcomes that are relevant for much longer periods.	Yes, in any circumstance we feel the difference between comparators is not (solely) brought about by changes in health (quality of life, survival) and (societal) costs.	Y	See earlier remarks.
s27	N	Patient preferences are important, but ideally they would be captured in a CUA that uses appropriate utility weights; also the CBA deals with the attributes of the test, not the consequences of testing (which is where benefits will accrue).	Perhaps in situations where there are two tests/services with different attributes, associated with similar health outcomes, WTP could identify which attributes (or combinations of attributes) are more highly valued by patients.	Y	A 'linked' approach is needed, because a genomic test on its own has no clinical value. But genomic sequencing cannot be addressed in an HTA as if it is a diagnostic test (i.e. false positives/negatives etc.): rather the 'performance' of genomic sequencing is more appropriately expressed in terms of the risk reduction associated with particular downstream interventions. Issues of costing could be addressed within current HTA frameworks.

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s28	U	I think it is useful for decision-makers to consider spillover effects and understand if there are any substantial costs (and benefits) imposed on other sectors from treatment adoption decisions. However, I'm not convinced presenting this information to them using a CBA type framework is the best way to do that.	See previous response. Yes. Especially when considering interventions that have clear inter-sectorial costs and benefits.	Y	The standard reference case approach needs to adapt to allow other measures to be used (beyond the EQ-5D) such as the capability approach. However, a balance needs to be sought between making analyses more technical and keeping them simple to use by decision-makers.
s29	Y	In my opinion, CBAs will be an important methodology in personalized medicine. This is because cost/QALY ICERs may not fully incorporate the value of knowing/value of knowledge. No it would not.	See previous response.	Y	The 'value of heterogeneity' and 'Expected Value of Individualized Care' framework merit further discussion, I think.
s30	N	Comparative judgements are the most important.	Probably not.	N	Many interventions have broad, hard to quantify implications.
s31	Y	I feel these methods should be used for national level decisions at least until genomic testing is more established in routine healthcare. We are doing similar comparisons of methodologies for regenerative medicine.	Of possible value in considering the longer term focus for programmes like the 100,000 Genomes Project, especially when applied to rare diseases.	Y	Beyond my expertise, but wider use in rare diseases (especially paediatric cases) should include a consideration of other (futile) diagnostic pathways - what patient groups call the "diagnostic odyssey".

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s32	U	No, I need a better understanding of the strengths and weaknesses of CBA approach.	Unable to answer.	N	I agree that HTA methods should be reviewed and adapted for the challenges of conducting economic evaluation of genetic or genomic tests. However, I believe these challenges apply to a range of different types of diagnostic tests and these are not exclusive problems for genetic/genomic based diagnostics.
s33	Y	I think it would be helpful to use CBA as well although it would mean that producing the evidence required by NICE etc. would become even more burdensome. CBA would be understood at that level but may not be understood at local level so it may need to be reconsidered (or provision made at the local level to make appropriate expertise available).	Not just in genomics but also public health and diagnostic - I think many of these same issues apply in complex technologies.	Y	I think yes but I am also worried that it could be a slippery slope. I would like to see the use of CBA in HTA but also there is a provision for genomics through companion diagnostics. I don't want genomic technologies to be unfairly valued but at the same time I think they need to be valued against non-genomic technologies in a clear and transparent way. I need to think on this a bit more...
s34	U	The value of the CBA seems to depend strongly on the quality of the DCE/WTP estimates. In this case, I would be concerned about the chosen target group (all CLL patients, not just those eligible for treatment) and low response rate/generalisability.	Yes.	Y	These challenges are likely to affect the assessment of a range of technologies, not just genetic/ genomic tests, hence development of methods to overcome these challenges seems useful.
s35	N	They represent alternative paradigms: there are overlaps, and double counting is possible.	No, but as part of a cost-consequence analysis, a comprehensive presentation of results would be helpful.	O	Not widespread adaptation because we need to be comparing like-analyses for alternative uses of health care resources. But, if it's broken then it needs to be fixed!

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s36	Y	CEAs which link data to outcomes are desirable but this data is not always available so CBAs may be more readily calculated. Commissioners report having a lack of health economic evidence so CBAs may be useful in this context or when directly comparing the costs of current tests.	CBA may be useful for comparing similar diagnostic tests but is probably less useful for introduction of novel far reaching technologies.	Y	A key issue is identifying the best measure for diagnostic tests and how to link these to outcomes - QALYs may not be best method here. Providing more economic evidence to support commissioners would be useful - if CBAs are more readily calculated this may help.
s37	N	The CBA performed via DCE only considers a narrow range of impacts, and was conducted in patients rather than general population.	Such analyses might inform patient co-pay or out of pocket cost policies.	O	In certain circumstances decisions will have to be made under conditions of greater uncertainty. Methods don't necessarily need to be different, but framework for decision-making could evolve to provide guidance in these circumstances, e.g. by considering quantitative evidence levels and the value of future research.

Resp	Should decision makers consider CBA results when deciding whether to adopt genomic testing in CLL? (Yes, No, Unsure) ^a	Please explain your response ^b	Are there any circumstances in which it would be helpful for healthcare decision-makers to consider CBA results when deciding whether an intervention should be adopted? ^c	Do you think that standard HTA methods should be adapted to reflect the challenges of conducting economic evaluations of genetic or genomic tests? (Yes, No, Only in specific circumstances)	If yes, please explain how you would like to see these methods adapted. Otherwise, please explain the reasoning behind your response ^d
s38	Y	Economic evaluation of NGS is challenging, as it is a diagnostic test, not always resulting in treatment implications. It might for example yield more diagnostics, and stop a diagnostic odyssey (and thereby prevent additional diagnostic testing). This benefit however, is difficult to capture in QALYs. Therefore a CBA taking into account other outcomes might be useful.	When a diagnostic test does not directly result in increased quality of life, but rather impacts non-health outcomes.	Y	Genetic and genomic tests are broadly applicable diagnostic tests, for very heterogeneous patient populations. Additionally, these results will often not have a direct impact on health-related quality of life. This does not mean that these tests don't have any value however.

CBA = cost-benefit analysis; CCA = cost-consequence analysis; CEA = cost-effectiveness analysis; CLL = chronic lymphocytic leukaemia; CUA = cost-utility analysis, DCE = discrete choice experiment; ENBS = expected net benefit of sampling; EVPI = expected value of perfect information; EVSI = expected value of sample information; HTA = health technology assessment; ICER = incremental cost-effectiveness ratio; LY = life year; NGS = next generation sequencing; NHS = national health service perspective; NICE = National Institute for Health and Care Excellence; QALY = quality-adjusted life year; Resp = respondent; WES = whole exome sequencing; WGS = whole genome sequencing; WTP = willingness-to-pay. ^a The exact wording for this question was: "In your opinion, would it be helpful for healthcare decision-makers at the national level (e.g. NICE, CADTH, PBAC) to consider cost-benefit analysis results alongside cost-effectiveness or cost-utility analysis results when determining whether or not to adopt genomic testing for CLL patients?". ^b The exact wording for this question was: "Please explain the reasoning behind your response to the previous question. Would you change your response if the question referred to healthcare decision-makers at a more local level (e.g. clinical commissioning groups in the UK NHS)?". ^c This question was worded differently depending on whether respondents thought that healthcare decision-makers should consider CBA results when determining whether to adopt genomic testing in CLL. Those who answered 'Yes' were asked: "Are there any other circumstances in which it would be helpful for healthcare decision-makers to consider cost-benefit analysis results alongside cost-effectiveness or cost-utility analysis results when deciding whether or not an intervention should be adopted in a healthcare system?". Those who were 'Unsure' or answered 'No' were asked: "Are there any circumstances in which it would be helpful for healthcare decision-makers to consider cost-benefit analysis results alongside cost-effectiveness or cost-utility analysis results when deciding whether or not an intervention should be adopted in a healthcare system?". ^d This question was worded differently depending on whether respondents thought that standard HTA methods should be adapted for genetic or genomic tests. Those who answered 'Yes' were asked: "Please explain how you would like to see these methods adapted". Those who responded 'Only in specific circumstances' or 'No' were asked: "Please explain the reasoning behind your response to the previous question".