Germline determinants of 5-fluorouracil drug toxicity and patient survival in colorectal cancer



Daniel Norris Rosmarin

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy, Clinical Pharmacology

Trinity Term 2013

Germline determinants of 5-fluorouracil drug toxicity and patient survival in colorectal cancer

Dan Rosmarin
Balliol College, University of Oxford, DPhil, Trinity 2013

Abstract Despite a decade of publications investigating the effect of germline polymorphisms on both toxicity related to treatment with 5-fluorouracil-based (5-FU) chemotherapy and prognosis following diagnosis with colorectal cancer (CRC), few genetic biomarkers have been identified convincingly. For 5-FU toxicity and CRC prognosis, in four results chapters, this thesis aims to validate previously-reported genetic biomarkers, identify new markers, determine the mechanistic basis of associated polymorphisms, and expand upon methods in the field.

The first three results chapters investigate genetic biomarkers for the prediction of toxicity caused by 5-FU-based treatment, particularly for the 5-FU prodrug capecitabine (Xeloda®, Roche). In the first, a systematic review and meta-analysis is performed for all variants that have been previously studied for an association with toxicity caused by any 5-FU-based drug regimen. 16 studies are analysed, including 36 previously-studied variants. Four variants show strong evidence of affecting a patient's risk of global (any) 5-FU-related toxicity upon analysis of both the existing data and over 900 patients from the QUASAR2 trial of capecitabine +/- bevacizumab (Avastin®, Roche/Genentech): *DPYD* 2846, *DPYD* *2A, *TYMS* 5'VNTR and *TYMS* 3'UTR. Next, 1,456 polymorphisms in 25 genes involved in the activation, action or degradation of 5-FU are investigated in 1,046 patients from QUASAR2. At a Bonferroni-corrected p-value threshold of 3.43e-05, three novel associations with capecitabine-related toxicity are identified in *DPYD* (rs12132152, rs7548189, A551T) and the previously-identified *TYMS* 5'VNTR and 3'UTR toxicity polymorphisms are refined to a tagging SNP (rs2612091) downstream of *TYMS* and intronic to the adjacent *ENOSF1*, the latter of which appears to be functional. Finally, a genome-wide investigation of 4.77 million directly genotyped or imputed SNPs identifies one variant (rs2093152 on chr20) as significantly associated with capecitabine-related diarrhoea ($p<5e-08$), though no associations meet this threshold for global toxicity.

In the study of CRC prognosis, a severe left truncation to the VICTOR trial is defined and shown to probably reduce statistical power but not bias effect estimates. Applying standard and novel genome-wide analysis approaches, a set of 43 SNPs are prioritised for future work.

With over one million new CRC cases annually, this work helps define biomarkers that could become broadly applicable in the clinical setting.

Acknowledgements

I have many thanks to give.

First, to Professor Tomlinson. Ian, you took in a student you did not admit then held nothing back in training me and overseeing my progress. You let me work on what interested me, even when other work seemed more urgent in your eye. You have taught me – and, more exceptionally, shown me – a new way of thinking, and I am thankful. I doubt many supervisors devote as much time and thought to their students as you do.

Second, to Dr Palles. Claire, you are a role model. Before Ian spent hours discussing results, you spent days showing me techniques. You would do this for me, you would do this for others in the lab and you would do this alongside all your own work. Your determination and work ethic are unmatched, but you always give time to others. Many students in the lab owe most of their skills to you.

To Dr Johnstone and Dr Midgley – Elaine and Rachel, thank you for bringing me to Oxford and identifying projects that would be of most interest to me. The course of my study was a surprise to us all and I am appreciative of your flexibility and unfaltering encouragement. Also, to Dr Jenny Taylor and the BRC group, whom I enjoyed collaborating with for research and from whom I appreciate financial support. Likewise, to Professor Len Seymour and the Departments of Clinical Pharmacology and Oncology, without which I would not have funded my DPhil. Of course, my greatest debt is owed to the Keasbey Memorial Foundation, whose trustees and supporters act generously to help students like me obtain an education in the UK.

To others in the Molecular and Population Genetics group – you have made this work even more enjoyable than it already is. Tim, Luke, Francesc, Michal, and others in the open plan, I appreciate your on-the-minute script help and laughs. David, for your help with the meta – even if you enjoyed it equally. Stefi, you made me feel welcome. To all others whose work I have built on, including Enric, Raj, Angela, Kimberley, and Axel. And to my collaborators, who suffered my onslaught of emails and requests.

To my family. Thank you for being so supportive, in all regards, when I am so far away.

Attributions

This thesis is a collection of the research carried out by the author in the Molecular and Population Genetics group at the Wellcome Trust Centre for Human Genetics, University of Oxford, as a member of the Department of Clinical Pharmacology (now Oncology). The work was performed between October 2010 and October 2013, supervised by Professor Ian Tomlinson, Dr Elaine Johnstone and Dr Rachel Midgley. The work contained in this thesis is original, though the work has been performed in collaboration with others, as has been cited and described in the text. No part of this research has previously been submitted for a degree at this or any other university.

The work described herein has formed the basis of two manuscripts and contributed to a third:

1. **Dan Rosmarin**, Claire Palles, David Church, Enric Domingo, Angela Jones, Elaine Johnstone, Haitao Wang, Sharon Love, Patrick Julier, Claire Scudder, George Nicholson, Anna Gonzalez-Neira, Miguel Martin, Daniel Sargent, Erin Green, Howard McLeod, Ulrich M Zanger, Matthias Schwab, Michael Braun, Matthew Seymour, Lindsay Thompson, Benjamin Lacas, Valerie Boige, Nuria Ribelles, Shoaib Afzal, Henrik Enghusen, Soren Astrup Jensen, Marie-Christine Etienne-Grimaldi, Gerard Milano, Mia Wadelius, Bengt Glimelius, Hans Garmo, Milena Gusella, Thierry Lecomte, Pierre Laurent-Puig, Eva Martinez-Balibrea, Rohini Sharma, Jesus Garcia-Foncillas, Zdenek Kleibl, Alain Morel, Jean-Pierre Pignon, Rachel Midgley, David Kerr, and Ian Tomlinson (2013)

Genetic markers of toxicity from capecitabine and other 5-fluorouracil-based regimens: investigation in the QUASAR2 study, systematic review and meta-analysis, *Journal of Clinical Oncology, in press.*

2. **Dan Rosmarin***, Claire Palles*, Alistair Pagnamenta, Kulvinder Kaur, Guillermo Pita, Miguel Martin, Enric Domingo, Angela Jones, Kimberley Howarth, Luke Freeman-Mills, Elaine Johnstone, Haitao Wang, Sharon Love, Claire Scudder, Patrick Julier, Ceres Fernandez-Rozadilla, Clara Ruiz-Ponte, Angel Carracedo, Sergi Castellvi-Bel, Antoni Castells, Anna Gonzalez-Neira, Rachel Midgley, Jenny Taylor, David Kerr, Ian PM Tomlinson (2013) A candidate gene study of capecitabine-related toxicity in colorectal cancer identifies new toxicity variants at DPYD and a putative role for ENOSF1 rather than TYMS, *manuscript submitted.*

3. David Church, Rachel Midgley, Enric Domingo, **Dan Rosmarin**, Kevin Maskell, Ian Tomlinson, David Kerr (2013) 'Toxgnostics': an Unmet Need in Cancer Medicine, *manuscript submitted.*

*These authors contributed equally to this work.

Contents

Abstract i

Acknowledgements ii

List of Figures xi

List of Tables xiii

List of Abbreviations xvi

1 Introduction 1

1.1 Background and Thesis Rationale 1

1.2 5-FU Toxicity 2

1.2.1 5-FU History and Schedules 2

1.2.2 Incidence of 5-FU Toxicity 3

1.2.3 Clinical Factors Affecting 5-FU Toxicity 6

1.2.4 Mechanisms of 5-FU Toxicity 11

1.2.5 Rationale for Predictive Biomarkers of 5-FU Toxicity 17

1.2.6 Genetic Biomarkers of 5-FU Toxicity 20

1.2.7 Non-genetic Biomarkers of 5-FU Toxicity 29

1.3 CRC Prognosis 29

1.3.1 Clinical Factors Affecting CRC Prognosis 30

1.3.2 Germline Genetic Biomarkers of CRC Prognosis 32

1.4 Approaches to Genetic Biomarker Analysis 34

1.4.1 Meta-analysis 34

1.4.2 Candidate-Based Analyses 37

1.4.3 Genome-Wide Association Studies 38
1.4.4 Survival Analysis 40
1.5 Thesis Outline 41

2 Materials and Methods 43
2.1 Samples Used 43
2.2 General Laboratory Methods 47
2.3 Roche/454 Next-Generation Sequencing 54
2.4 Imputation and Association Testing of Imputed Data 69
2.5 Statistical Methods 72
2.6 Results Interpretation Tools 84

3 Systematic Review and Meta-Analysis of Genetic Markers of 5-FU Toxicity and Investigation in QUASAR2 86
3.1 Introduction, Aims and Overview 86
3.2 Challenges Particular to this Analysis 88
3.3 Literature Search 89
3.4 Study Filtering and Formation of Patient Sets 90
3.5 Identification of Included Variants and Aggregation of Included Data 93
3.6 Toxicity Incidence by Regimen and Formation of a "Global" Measure 96
3.7 Assessment of Publication Bias 101
3.8 Summary of Analysis Plan 102
3.9 QUASAR2 Results for Previously Studied Variants and Development of the *TYMS* Score Test 104
3.10 Capecitabine Monotherapy Meta-Analysis Results 108
3.11 Infusional and Bolus 5-FU Monotherapy Meta-Analysis Results . . . 110
3.12 Combination Therapy 5-FU Meta-Analysis Results 113
3.13 DPYD Deficiency Syndrome Literature Review and *DPYD* Combined Analysis 115
3.14 Summary of Significant Findings 119
3.15 Consideration of Further Pooling Across 5-FU Regimens 128
3.16 Review of Commercially Sold 5-FU Toxicity Predictive Kits and ROC Analysis in QUASAR2 131

3.17 Discussion 133

4 Candidate Gene Analysis of Capecitabine-Related Toxicity 138

4.1 Introduction 138
4.1.1 Background 138
4.1.2 Aims and Overview 140
4.1.3 Patient Inclusion 141
4.1.4 Toxicity Data 142
4.2 Imputation of SNP Array Data – Strategy, Method and Technical Results 145
4.2.1 Array Pooling Strategy Options 146
4.2.2 SNPTEST Downstream Filtering 147
4.2.3 SNPTEST Upstream Filtering 148
4.2.4 Pooling Strategy Comparison 149
4.2.5 Summary of Imputation Method Used 149
4.3 Exome Array Data Quality Control and Genotype Calling – Technical Results 152
4.4 Exon Sequencing Analysis – Technical Results 157
4.5 5-FU/Capecitabine Candidate Gene Analysis: Toxicity Variant Identification Results 163
4.5.1 Overview of Analysis Method and Included Data 163
4.5.2 Overview of SNP and Exome Array Results 165
4.5.3 *DPYD* rs12132152 is Associated with Global Capecitabine Toxicity, Driven by HFS 165
4.5.4 *DPYD* rs7548189 is Associated with Global Capecitabine Toxicity, Driven by Multiple Phenotypes 170
4.5.5 *TYMS/ENOSF1* rs2612091 is Associated with Global Capecitabine Toxicity, Driven by HFS 174
4.5.6 rs2612091 Impacts *ENOSF1*, not *TYMS*, at the mRNA Expression Level 176
4.5.7 rs2612091 Better Explains Toxicity Variation than do the *TYMS* 5'VNTR and/or 3'UTR Polymorphisms 177
4.5.8 Set-Based Tests of Capecitabine/5-FU Pathway Genes 181

4.5.9 *DPYD* A551T is a Novel Rare 5-FU/Capecitabine Toxicity Variant . . . 182
4.5.10 ROC Analysis of New and Known 5-FU/Capecitabine Toxicity Variants in QUASAR2 . . . 186
4.6 Discussion . . . 187

5 Capecitabine Toxicity GWAS of QUASAR2 and SCP Studies 190
5.1 Introduction . . . 190
5.1.1 Background . . . 190
5.1.2 Aims and Overview . . . 191
5.2 Heritability of Capecitabine-Related Toxicity . . . 192
5.3 GWAS of Capecitabine-Related Toxicity in the QUASAR2 Study . . 194
5.4 Genome-Wide Meta-Analysis of Capecitabine-Related Toxicity in the QUASAR2 and SCP Studies . . . 197
5.5 Correlation Among Capecitabine-Related Toxicity Phenotypes in QUASAR2206
5.6 Discussion . . . 210

6 Germline Markers of Prognosis in CRC 212
6.1 Introduction . . . 212
6.2 Aim and Overview . . . 214
6.3 Outcome Measures . . . 215
6.4 VICTOR Dataset Summary and Truncation . . . 216
6.4.1 VICTOR Clinical and Molecular Data . . . 216
6.4.2 Left Truncation of the VICTOR Dataset . . . 217
6.4.3 Blood Collection Lag and VICTOR Prognosis . . . 220
6.4.4 Covariate Selection . . . 223
6.5 Heritability of CRC Prognosis . . . 226
6.6 Prognosis in VICTOR by Cox Proportional Hazards Method . . . 228
6.7 Prognosis Based on Blood Collection Lag . . . 234
6.7.1 Prognosis Based on Blood Collection Lag: GWAS Within VICTOR . . . 235
6.7.2 Prognosis Based on Blood Collection Lag: GWAS Within CORGI238
6.8 Prognosis Based on VICTOR/QUASAR2 MAF Differences . . . 239

6.9 Comparison of Results Across Methods 242
6.10 Novel Method for Survival Analyses of Truncated Datasets 244
6.11 Discussion 248

7 Discussion **251**

Appendices **258**

A Power Calculations for 5-FU/Capecitabine Association Studies **258**

B 5-FU Toxicity Meta-Analysis QUASAR2 Genotyping **261**

C Literature Search Strategy for 5-FU Toxicity Meta-Analysis **264**

D 5-FU Toxicity Meta-Analysis 2-by-2 Table Counts of Included Data **267**

References **276**

List of Figures

1.1 Mechanisms of 5-FU Activation, Action and Catabolism 12
1.2 Commercial Kits for Prediction of 5-FU Toxicity 29

2.1 Roche/454 Amplicon Sequencing Patient Pooling 56
2.2 Roche/454 Amplicon Sequencing Workflow 57
2.3 Roche/454 Amplicon Sequencing Primer Design 59
2.4 Sequencing Homopolymer Example 68

3.1 Meta-Analysis Study Filtering 91
3.2 Meta-Analysis Funnel Plots 102
3.3 *TYMS* Forest Plot for FOLFOX 115
3.4 *TYMS* 5'VNTR Forest Plot 121
3.5 *TYMS* 3'UTR Forest Plot 122
3.6 *DPYD* *2A Forest Plot 123
3.7 *DPYD* 2846 Forest Plot 124
3.8 *MTHFR* 1298 Forest Plot 125
3.9 Consistency of Effect Estimates Across Regimens 130
3.10 ROC Analysis in QUASAR2 of Meta-Analysis Variants 133

4.1 Imputation Array Pooling Options 152
4.2 QUASAR2 Exome Array PCA 154
4.3 Exome Array Heterozygosity Rates with Genome Studio and zCall . . 156
4.4 Bioanalyzer Results for Roche/454 Pool 1 159
4.5 Roche/454 Final Coverage Achieved 161
4.6 Identification of True Variants from Roche/454 Sequencing 162

4.7 LocusZoom Plots for rs12132152 167
4.8 LocusZoom Plots for rs7548189 171
4.9 LocusZoom Plots for rs2612091 175
4.10 *TYMS* and *ENOSF1* DNA Coding Overlap 176
4.11 *TYMS* and rs2612091 LD Structure 178
4.12 Identification of New Putative Toxicity Variants from Roche/454 Sequencing 183
4.13 ROC Analysis Following Candidate Gene Study 187

5.1 PCA of 940 QUASAR2 Patients 193
5.2 Manhattan and QQ Plots of Capecitabine GWAS for all Phenotypes . 200
5.3 HFS and Diarrhoea Effect Size Correlation 208

6.1 VICTOR Patient Flow Diagram 217
6.2 Genetic Left Truncation in VICTOR 219
6.3 VICTOR OS and RFS Kaplan Meier Curves 222
6.4 VICTOR OS Schoenfeld Residuals 225
6.5 PCA of 918 VICTOR Patients 227
6.6 VICTOR Prognosis Cox PH Manhattans 231
6.7 VICTOR Prognosis Cox PH QQ Plots 232
6.8 VICTOR Prognosis Cox PH OS and RFS Beta Comparison 233
6.9 VICTOR Blood Lag Period GWAS 237
6.10 VICTOR/QUASAR2 MAF Comparison 241
6.11 VICTOR Full Model Results 247

List of Tables

1.1 Description of Common 5-FU-Related Toxicity Phenotypes 4
1.2 CTCAEv4 Adverse Event Grades 4
1.3 QUASAR2 Dose Reduction and Modification 5
1.4 5-FU Candidate Gene Names and Abbreviations 13
1.5 Previously Reported SNP Associations with 5-FU Toxicity 21
1.6 Functional Evidence for 5-FU Toxicity Variants 26
1.7 CRC 5-Year Overall Survival by Stage 31
1.8 Examples of Meta-Analysis 37

2.1 PCR Reagents 49
2.2 PCR Troubleshooting 50
2.3 KASPar Reagents 51
2.4 Sanger Sequencing Reagents 52
2.5 Roche/454 Amplicon Sequencing *DPYD* Primers 61
2.6 Roche/454 Amplicon Sequencing *TYMS* and *UPP2* Primers 62
2.7 Genetic Models 76
2.8 Results Interpretation Tools 85

3.1 Summary of Patient Sets for Meta-Analysis 94
3.2 Aggregated Variants and Studies for Meta-Analysis 97
3.3 Excluded Variants for Meta-Analysis 98
3.4 Individual Toxicity Phenotype Incidence by 5-FU Regimen 100
3.5 Global Toxicity Incidence by 5-FU Regimen 101
3.6 QUASAR2 Toxicity Results for Previously Studied Variants 106
3.7 Capecitabine Monotherapy Meta-Analysis Results 109

3.8 Infusional/Bolus Monotherapy Meta-Analysis Results 111
3.9 Combination Therapy Meta-Analysis Results 114
3.10 DPYD Deficiency Syndrome Variants 116
3.11 *DPYD* Rare Variant Group Test 118
3.12 *DPYD* Rare Allele and *TYMS* Score Test Results 127
3.13 Commercial Kits for Prediction of 5-FU Toxicity 132

4.1 QUASAR2 Clinical Characteristics for Candidate and GWAS Analysis 143
4.2 QUASAR2, SCP and EPICOLON Toxicity Data 144
4.3 Imputation Downstream Filtering 148
4.4 Comparison of IMPUTE2 and SNPTEST2 Info Scores 150
4.5 Inflation Arising from Array Pooling Options 151
4.6 Roche/454 Sequencing Round 1 – Low Coverage Amplicons 159
4.7 Number of Variants Tested Per Candidate Gene 164
4.8 rs12132152 Associations with Toxicity 168
4.9 rs7548189 Associations with Toxicity 172
4.10 Toxicity Phenotype Interactions for the rs7548189 SNP Associations . 173
4.11 rs2612091 Associations with Toxicity 174
4.12 AIC for rs2612091 and the *TYMS* UTR Polymorphisms 179
4.13 Independent Effect Analysis of the rs2612091 and *TYMS* 5'VNTR and 3'UTR Polymorphisms 180
4.14 5-FU/Capecitabine Pathway Set-BasedTest Results 182
4.15 Novel Putative Toxicity Variants from Roche/454 Sequencing 185

5.1 Heritability of Capecitabine Toxicity 194
5.2 QUASAR2-Only GWAS of Capecitabine-Related Toxicity 196
5.3 QUASAR2/SCP GWAS of Capecitabine-Related Toxicity 199
5.4 Correlation Among Toxicity Phenotypes 208
5.5 Independent Effect of SNPs Associated with Multiple Phenotypes . . 209

6.1 VICTOR Covariate Selection 224
6.2 Heritability of CRC Prognosis 228
6.3 Comparison of CRC Prognosis Model Outputs 243

A.1 5-FU Toxicity Power Calculations – 90% Power with OR=1.5 259
A.2 5-FU Toxicity Power Calculations – 90% Power with OR=3.0 260

B.1 5-FU Toxicity Meta-Analysis QUASAR2 Genotyping 262

C.1 5-FU Toxicity Meta-Analysis Literature Search 265

D.1 5-FU Toxicity Meta-Analysis 2-by-2 Table Counts of Included Data . 268

List of Abbreviations

5-FU	5-fluorouracil
AIC	Akaike information criterion
AJCC	American Joint Committee on Cancer
AUC	area under curve (eg, under the ROC curve)
BIC	Bayesian information criterion
BOL	bolus 5-FU
CAP	capecitabine 5-FU
CEU	Utah residents with ancestry from northern and western Europe (aka CEPH)
CI	confidence interval
CRC	colorectal cancer
DEPC	diethylpyrocarbonate
DFS	disease-free survival
DMSO	dimethyl sulfoxide
dNTPs	deoxynucleotide triphosphates
FDR	false discovery rate
FOLFIRI	infusional 5-FU, folinic acid and irinotecan
FOLFOX	infusional 5-FU, folinic acid and oxaliplatin
GWAS	genome-wide association study
HFS	handfoot syndrome

INF	infusional 5-FU
IS	info score
LD	linkage disequilibrium
MAF	minor allele frequency
mCRC	metastatic colorectal cancer
NPV	negative predictive value
OS	overall survival
PCA	principal component analysis
PCR	polymerase chain reaction
PPV	positive predictive value
PRISMA	preferred reporting items for systematic reviews and meta-analyses
RFS	recurrence-free survival
ROC	receiver operating characteristic
SCP	Spanish Capecitabine Pharmacogenetic study
SNP	single (simple) nucleotide polymorphism
UTR	untranslated region
VNTR	variable number tandem repeat
XELOX	capecitabine and oxaliplatin

Chapter 1

Introduction

1.1 Background and Thesis Rationale

Colorectal cancer (CRC) is an epithelial malignancy of the colon or rectum. In terms of genomic alteration, tumours at these locations are largely similar (29), though treatment recommendations do vary (99; 100) and molecular features either change or shift when comparing tumours of the right (proximal colon) and left (distal colon and rectal) side (217). However, for many aspects of treatment and in many clinical trials, the 70% of CRC patients with colonic tumours and the 30% of CRC patients with rectal tumours are considered in a single group, and such practice is followed throughout this thesis, using the single classification of CRC unless otherwise noted.

Globally, there are an estimated 1.2 million new CRC cases annually, with a far higher incidence in developed countries, suggesting an environmental influence (103). In addition, CRC has a genetic component, including high penetrance syndromes such as familial adenomatous polyposis (FAP) and lower penetrance predisposition among sporadic CRC, accounted for by twin studies that suggest a 35% heritable component to the disease risk (123).

Among the surgical, radio- and chemotherapy options used to treat CRC pa-

tients, outcome remains modest, with 5-year survival of 83% among patients with localised disease (stage II), 60% among patients with regionalised disease (stage III) and 8% among patients with metastatic disease (mCRC; stage IV) (147). Treatment options all come at a cost to patients. The tradeoffs of treatment are particularly balanced for patients undergoing 5-fluorouracil-based (5-FU) chemotherapy, as 10-30% of patients suffer drug toxicity (Chapter 3) in exchange for a 5-20% stage-dependent survival benefit in the adjuvant setting (75; 171) and first-year-only survival benefit in the metastatic setting (110). Predicting which patients can be safely treated by chemotherapy, and the appropriate dose for doing so, is thus crucial in determining who stands to benefit from chemotherapy. Similarly, accurately predicting the survival of a newly diagnosed patient (ie, their prognosis) helps guide treatment decisions around the risk/benefit equilibrium.

Despite over a decade of publications in the field, genetic markers predicting toxic response to 5-FU-based chemotherapy and patient prognosis remain poorly characterised. Through both new investigation into germline genetic markers of these outcomes and a review and meta-analysis of previously reported investigations, this thesis aims to improve our ability to personalise colorectal cancer treatment decisions.

1.2 5-FU Toxicity

1.2.1 5-FU History and Schedules

5-fluorouracil (5-FU) is a chemotherapeutic drug introduced in the late 1950s. As a non-targeted agent, 5-FU indiscriminately affects cancer and non-cancer cells, inhibiting DNA synthesis and both DNA and RNA function in all exposed cells that are actively creating genetic material (76). However, elevated cellular growth and altered levels of various enzymes mean that 5-FU is particularly potent in causing tumour cell death (full details in Section 1.2.4).

5-FU is administered in one of three regimens (ie, schedules), used either as monotherapy or as the backbone for combination therapy regimens. The original schedule is bolus 5-FU (BOL), in which a single intravenous dose is injected on a daily or twice-daily basis. Beginning in the 1990s, the chemosensitizer leucovorin (ie, folinic acid) was added to bolus 5-FU as part of the Mayo and Roswell Park regimens (148; 213). Subsequently, due to improvements in toxicity and efficacy, bolus 5-FU was largely replaced with infusional 5-FU (INF), in which 5-FU is administered intravenously in a continuous fashion, often as the de Gramont regimen (45). Infusional 5-FU forms the backbone of two commonly used combination therapy regimens, FOLFOX (INF + OX, oxaliplatin) and FOLFIRI (INF + IR, irinotecan), which are first- or second-line therapies for both metastatic and non-metastatic CRC (46; 58). In the years approaching the turn of the millennium, an oral prodrug of 5-FU was developed, capecitabine (CAP, trademark: Xeloda®, Roche), and this drug has grown in popularity since the early 2000s, both as monotherapy and combination therapy (eg, CAP + OX, "XELOX") (32; 199). Capecitabine is administered orally, after which it is converted intracellularly into 5-FU, with specificity gained through preferential activation in tumour cells (144).

1.2.2 Incidence of 5-FU Toxicity

Due to the indiscriminate nature of 5-FU-based drugs, 5-FU toxicity is common (10%-30%, depending on regimen), including gastrointestinal toxicities (diarrhoea and nausea/vomiting), haematologic toxicities (neutropaenia and thrombocytopaenia) and others. Table 1.1 defines the 5-FU-related toxicity phenotypes referred to throughout this thesis.

For patients undergoing treatment, these individual toxicities are quantified and tracked using the World Health Organization (WHO) or Common Terminology Criteria for Adverse Events (CTCAE) grading scales, the latter of which is current best

Table 1.1: Description of the toxicity phenotypes occurring with at least 5% incidence following treatment with 5-FU-based regimens.

Toxicity	Desription
Diarrhoea	Increase in stools per day relative to baseline
Mucositis/stomatitis	Ulceration of mucous membranes, including mouth
Neutropaenia	Reduction in neutrophils (ie, white blood cells)
Thrombocytopaenia	Reduction in platelets (aka myelosuppression)
Nausea/vomiting	Appetite loss and/or any emesis
Handfoot syndrome	Redness, swelling, discomfort, and/or skin peeling of hands and feet (aka palmar-plantar erythrodysesthesia syndrome or "HFS")

Table 1.2: Categorisation of adverse events (AE) into toxicity grades based on CTCAEv4, and common impact on treatment decisions for patients given 5-FU-based regimens (202).

Grade	Severity	Treatment impact
1	Mild	Clinical or diagnostic observation only
2	Moderate	Non-invasive intervention only
3	Severe	Hospitalisation and 25% dose reduction
4	Life threatening	Urgent intervention and at least 50% dose reduction
5	Death from AE	

practice. On the CTCAE scale, grade 0 equates to non-existent adverse events, grade 1 or 2 events are moderate but generally do not influence treatment decisions unless recurrent or precursors of life threatening toxicities, and grade 3 or 4 events are "dose limiting," meaning that a patient's course of treatment is altered with a delay and/or dose reduction (202). Table 1.2 shows the definitions of CTCAE toxicity grades, according to the most recent release: version 4, released in 2009 and modified in 2010 (Institute). The dose and schedule reduction protocol of the QUASAR2 trial of capecitabine with or without bevacizumab, which features in this thesis, is summarised in Table 1.3.

Despite interventions, onset of toxicity can be rapid and severe, resulting in ap-

Table 1.3: Summary of the QUASAR2 dose reduction and schedule modification protocol, modified from the trial protocol. Haematologic toxicities, namely neutropaenia, are considered most life threatening, though incidence of these phenotypes is lower for patients given capecitabine relative to other forms of 5-FU. Starting dose for most patients is 1250 mg/m^2. Events are categorised according to CTCAEv3.

	Worst event	**Treatment delay**	**Dose reduction**
Diarrhoea	grade 2	Wait until recovery to grade 0 or 1	0%
	grade 3		25%
	grade 4		50%
Mucositis/stomatitis	grade 2	Wait until recovery to grade 0 or 1	0%
	grade 3		25%
	grade 4		50%
Neutropaenia	grade 1	Wait until recovery to grade 0	Discontinue if no recovery in 3 weeks
Thrombocytopaenia	grade 1	Wait until recovery to grade 0	Discontinue if no recovery in 3 weeks
Nausea/vomiting	grade 3/4	Wait until recovery to grade 0 or 1	Discuss with trial coordinator
Handfoot syndrome	first grade 2	Wait until recovery to grade 0 or 1	0%
	second grade 2		25%
	first grade 3		50%
	second grade 3		further 25-50%

proximately 0.5-1% drug-induced mortality (ie, grade 5) for monotherapy regimens, and up to 3% drug-induced mortality for combinations regimens (169). Notably, drug-induced mortality does not tend to result from a single given toxicity phenotype, but rather death results from a "cluster" of toxicities occurring simultaneously or in rapid succession (165). Such apparent correlation among phenotypes suggests that the use of a "global" 5-FU-related toxicity measure is clinically relevant, given that a patient who experiences one type of severe toxicity is susceptible to others, at least among patients experiencing the most severe toxicity. If clinically relevant, use of a global measure would also aid association studies, as it improves statistical power through higher event rates. Use of this measure is further discussed in Chapter 3.

1.2.3 Clinical Factors Affecting 5-FU Toxicity

Inter-patient differences in 5-FU toxicity are partly the result of clinical factors, including those related to the particular 5-FU regimen administered and those related to patient characteristics. For the purposes of this thesis, "clinical factors" refers to those other than germline genetics.

While it is well understood by clinicians that incidence of adverse events (ie, drug toxicity) resulting from treatment with 5-FU varies by regimen, this understanding is based largely on results from individual phase II/III trials and the ranges have been poorly summarised in the literature. One result produced by the systematic review described in Chapter 3 of this thesis is a summary of dose limiting 5-FU toxicity incidence across 5-FU regimens.

Toxicity events differ in the cycle of treatment at which they tend to occur. For example, handfoot syndrome has been observed at higher incidence in the later cycles of treatment as compared to in the initial cycle (115). As such, early toxicity events can give some sense of toxicity events that should be expected if the course of treatment is not changed. Following from this, it is common clinical practice to implement a dose

reduction. In the X-ACT trial comparing capecitabine to infusional 5-FU/leucovorin, for example, 42% of capecitabine patients and 44% of 5-FU/LV patients required a dose reduction (200). This is just a subset of the 57% of capecitabine patients and 52% of 5-FU/LV patients requiring some form of dose modification in that trial.

The particular treatment regimen given to a patient can affect the timing of toxicity (ie, which cycle of treatment). Again looking at the X-ACT trial of adjuvant CRC treatment, capecitabine patients were able to remain at full dose of capecitabine longer than were 5-FU/LV patients (median of 78 days vs 41 days) (200). Considering the PETACC-3 trial of adjuvant infusional 5-FU/LV with or without irinotecan, patients given somewhat higher doses of 5-FU as part of a monotherapy regimen experienced higher rates of drug discontinuation due to adverse events than did their counterparts given a combination regimen (201). This was despite the use of irinotecan in the latter group, which led to overall higher rates of grade 3/4 adverse events. Generally, the cycle of treatment at which toxicity occurs has been sparsely reported and/or summarised in results of clinical trials. This is unfortunate, given that the absolute occurrence of high toxicity does not differentiate between patients who should start treatment at full dose and subsequently be dose reduced and patients who should be given a reduced dose from the start.

Patient characteristics, particularly age and gender, consistently appear as independent predictors of 5-FU-related toxicity. The effects of these factors were first identified by Stein et al using multivariate and univariate analysis of patients stratified into two age groups (above and below 70 years of age) and the two sexes (186). In this study, males and younger patients were independently shown to exhibit lower rates of "any severe" 5-FU-related toxicity (sex: 30% vs 56%, $p<0.0001$; age: 36% vs 58%, $p<0.001$). These two clinical factors remained significant in multivariate logistic regression, in which the particular 5-FU-based regimen used was similarly significant. As expected, the highest incidence of "any severe" toxicity was experienced by elderly

women (81%). Treatment length and patient performance status were not significant predictors of toxicity.

Interestingly, in the Stein et al study, age and sex not only influenced the overall incidence of dose limiting toxicity, but they shifted the toxicity profile experienced by patients. While the general population experienced a high incidence of diarrhoea, which is considered a "clinically manageable" toxicity, the more elderly patients suffered higher rates of leukopaenia (of which neutropaenia is a component) – considered among the most "clinically risky" 5-FU-related toxicities (24% vs 10%, $p<0.005$). Females, likewise, experienced a more than two-fold increase in this toxicity phenotype (21% vs 7%).

The effects of age and gender have been confirmed in subsequent studies, and the causes have been further refined. Cassidy et al confirmed that both age and gender affect capecitabine toxicity (33). This report additionally showed that age is not a causal predictor of toxicity, but rather that age is negatively correlated with renal function. As such, age was a non-significant predictor of toxicity when adjusting for creatine clearance. Interestingly, impaired renal function was shown to increase levels of 5'dFUR (Section 1.2.4). This capecitabine metabolite is directly converted to 5-FU by the enzymes *UPP1* and *UPP2*, both in the tumour, where the drug takes on its desired role, and in the liver, where the drug is quickly metabolised (125; 163). However, these enzymes also have high expression in the kidneys, and thus high renal 5'dFUR levels may cause toxicity through activation to 5-FU outside of the tumour or liver. Regarding gender, reports have shown that low lean body mass is associated with higher toxicity levels (157), which, given that 5-FU is dosed according to body surface area, could explain both why women and the elderly experience more toxicity than young men. Interestingly, women experience similar survival, suggesting that women and their tumours are not inherently more sensitive to the drug at the bioavailable levels (130; 181).

Zalcberg et al confirmed the impact of age and gender and further showed that the incidence of grade 3/4 toxicity depends on cycle number, showing that incidence fell from 39% to as low as 12% between the first and subsequent 5-FU treatment cycles (219). This analysis did not, however, account for the fact that patients with toxicity events would be less likely to continue their course of treatment. Although historically age and renal impairment might have been important causes of toxicity in treatment populations, these factors are now well recognized, and thus adjusted or accounted for. In the QUASAR2 trial of capecitabine with and without bevacizumab discussed in this thesis, for example, patients over 70 years of age or with evidence of renal impairment began treatment with a reduced dose.

Regional differences in FU-related toxicity also exist. Comparing toxicity rates among patients in the US, East Asia and the "rest of world" (mostly Europe), collected from three trials, Haller et al showed that US patients experience higher rates of grade 3/4 events (RR=1.77, $p<0.001$), dose reduction (RR=1.72, $p<0.001$) and drug discontinuation (RR=1.83, $p<0.001$) (80). The authors of this study named many possible causes of these regional differences, including reporting bias, dietary folate intake and cultural differences that affect a patient's will to continue treatment. Importantly, the authors did not adjust for or stratify analyses based on ethnicity, meaning that regional differences may be acting as proxies for underlying genetic differences. The authors argue that the ethnic diversity of the US leads to genetic diversity that makes it unlikely for a commonly held set of alleles to be increasing incidence of toxicity; they fail to consider, however, that a certain US ethnic group within the study could be driving the effect. Clinically, as a result of regional differences, capecitabine dosing guidelines recommend lower doses in the US than in Europe (1000 mg/m^2 vs 1250 mg/m^2).

No publication has assessed the effect of ethnicity (or region adjusted for ethnicity) on 5-FU-related toxicity; given that some studies exist for the effect of ethnicity on

outcome (ie, survival), this literature gap is surprising. However, it has been shown that Caucasian and Japanese patients do not exhibit differences in capecitabine pharmacokinetics, as measured by levels of various 5-FU metabolites (161). The only metabolite level that differed between the populations was that of higher levels of FBAL, the end product of 5-FU catabolism, among Caucasians. The authors discredit FBAL levels as having no documented impact on toxicity or outcome; plausibly, however, the higher levels of FBAL might suggest that Caucasians depend more heavily on the DPYD pathway for drug catabolism, while Japanese may have an alternative degradation pathway (see Figure 1.1 for further details on 5-FU metabolism pathways). For studies of capecitabine, higher FBAL levels may indicate that a larger proportion of the prodrug is converted to 5-FU (and subsequently catabolised fully).

One non-genetic cause of regional differences in 5-FU toxicity may be diet, particularly the amount of folate in one's diet, given that reduced folate is necessary for 5-FU action. Rats given a high folate diet have been shown to suffer from higher levels of toxicity (24). These results have been confirmed for patients given capecitabine, showing a positive correlation between pretreatment serum levels and toxicity at any cycle of treatment, though sample size, effect size and significance levels were all small (n=60, p=0.04, 95% CI 1.0-1.2) (176). As pointed out by Midgley et al, the US policy of fortifying food with folinic acid may account for the regional differences in toxicity (141); however, given the effect size of folate, this factor seems unlikely to be the only cause.

Unfortunately, many reports of biomarkers of 5-FU toxicity fail to include any of these clinical factors as covariates (or, it turns out, even collect data on these factors), hindering adjusted analyses when pooling data across studies.

1.2.4 Mechanisms of 5-FU Toxicity

5-FU can be administered either directly or indirectly. Direct administration of 5-FU takes the form of an injection, either as a single bolus or via slow infusion of the 5-fluorouracil molecule. Indirect administration of 5-FU occurs with the oral prodrug capecitabine (Xeloda®, Roche), which is absorbed in the gut and subsequently converted to 5-FU. Other oral prodrugs are in development but have yet to be approved. As described in Section 1.2.2, 5-FU-related toxicity results from treatment via all administration routes.

Understanding the sources of 5-FU-related toxicity first requires an understanding of the mechanisms of 5-FU activation, drug action and catabolism. This is because inter-individual variation in these processes, at least somewhat due to enzyme activity, could affect one's predisposition to toxicity. Figure 1.1 shows (a) the common drug activation pathway and mechanism of action, (b) alternate drug activation pathways and mechanisms of action and (c) the 5-FU catabolism pathway by which the drug is degraded. Given the genetic biomarker focus of this thesis, the figure shows activation, action and catabolism processes according to their various enzymes, providing a list of candidate genes for the study of 5-FU toxicity, as performed in Chapter 4. For reference, all enzyme abbreviations and protein names are defined in Table 1.4.

Prior to discussing how certain enzymes predispose patients to 5-FU toxicity, it is useful to have a broad understanding of 5-FU pharmacokinetics and pharmacodynamics. 5-FU is able to exert a cytotoxic effect only once it is converted to one or more of its active forms: FdUMP, FdUTP and FUTP. This activation occurs preferentially within tumour cells due to elevated concentration of key enzymes. For capecitabine, it is first necessary to convert the inactive prodrug to 5-FU (also not directly cytotoxic), largely within the liver, and then to convert 5-FU to these active forms. Prior to activation, much of the drug is rapidly broken down within the liver. 5-FU-related toxicity occurs when non-tumour cells are exposed to high con-

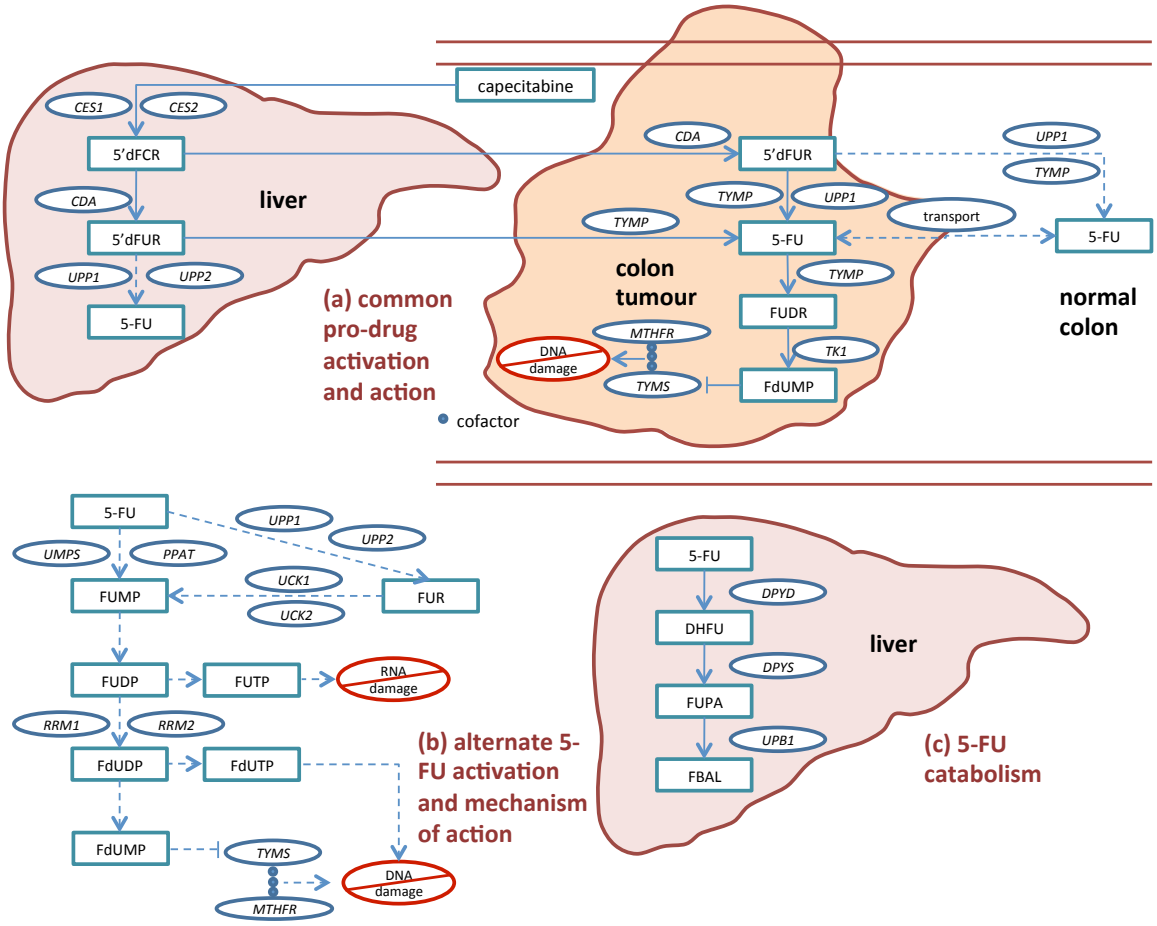


Figure 1.1: Capecitabine is an oral 5-FU prodrug that is rationally designed so that concentrations of the cytotoxic metabolites FdUMP, FdUTP and FUTP are higher within malignant cells than within normal cells. After absorption in the gut, most of the drug activation occurs via the common prodrug activation route (a), starting in the liver and finishing in the tumour. Additionally, 5-FU can be converted to its active compounds via alternate activation routes (b) in both colon tumour cells and cells from multiple other tissues. Toxicity may occur if non-target tissue is exposed to activated capecitabine/5-FU (ie, FdUMP, FdUTP and FUTP), either when 5-FU is activated prior to arrival at the tumour or when it exits target tissue and is subsequently activated. 5-FU released by cells into circulation may be quickly metabolised (c) in the liver by enzymes including *DPYD*. **Abbreviations**: mechanism of action (MOA), 5-fluorouracil (5-FU), fluorodeoxyuridine monophosphate (FdUMP), fluorodeoxyuridine diphosphate (FdUDP), fluorodeoxyuridine triphosphate (FdUTP), fluorouridine triphosphate (FUTP), 5'-deoxy-5'-fluorocytidine (5'dFCR), 5-fluorourodeoxyuridine (5'dFUR), fluorodeoxyuridine (FUDR), fluorouridine monophosphate (FUMP), fluorouridine (FUR), fluorouridine diphosphate (FUDP), fluorodeoxyuridine monophosphate (FdUMP), dihydrofluorouracil (DHFU), alpha-fluoro-beta-ureidopropionic acid (FUPA), alpha-fluoro-beta-alanine (FBAL). Largely adapted from four sources (51; 127; 194; 211).

Gene symbol	Gene name
ABCB1	ATP-binding cassette, sub-family B
ABCC3	ATP-binding cassette, sub-family C, member 3
ABCC4	ATP-binding cassette, sub-family C, member 4
ABCC5	ATP-binding cassette, sub-family C, member 5
ABCG2	ATP-binding cassette, sub-family G, member 2
CDA	cytidine deaminase
CES1	carboxylesterase 1 isoform a precursor
CES2	carboxylesterase 2 isoform a precursor
DPYD	dihydropyrimidine dehydrogenase
DPYS	dihydropyrimidinase
MTHFR	methylenetetrahydrofolate reductase
PPAT	phosphoribosyl pyrophosphate amidotransferase
RRM1	ribonucleoside-diphosphate reductase subunit 1
RRM2	ribonucleoside-diphosphate reductase subunit 2
SLC22A7	solute carrier family 22 member 7 isoform b
SLC29A1	equilibrative nucleoside transporter 1
TK1	thymidine kinase 1
TYMP	thymidine phosphorylase
TYMS	thymidylate synthetase
UCK1	uridine-cytidine kinase 1 isoform a
UCK2	uridine-cytidine kinase 2 isoform a
UMPS	uridine monophosphate synthase
UPB1	beta-ureidopropionase
UPP1	uridine phosphorylase 1
UPP2	uridine phosphorylase 2

Table 1.4: Names and gene symbols for the 25 enzymes involved in 5-FU/capecitabine activation, drug action and catabolism.

centrations of the drug's active forms. The apparent causes of 5-FU-related toxicity are similar for both injected and oral 5-FU, except that additional sources of toxicity exist during the activation of capecitabine to 5-FU. However, these additional sources of 5-FU toxicity are compensated for by the fact that capecitabine activation is tumour-specific, and thus the drug's toxicity profile is generally less severe than that of direct 5-FU (199).

Based on mechanism, germline genetic variants in genes that code for the enzymes in the capecitabine/5-FU pathway may cause predisposition to toxicity in one of four ways:

1. uptake speed and conversion of the oral prodrug capecitabine to 5-FU prior to arrival at the tumour;

2. leakage of 5-FU and other pre-5-FU capecitabine metabolites from the tumour into the surrounding normal tissue;

3. conversion of 5-FU into its active forms outside the tumour, via alternative activation pathways; and

4. reduced rates of drug degradation.

These four causes of 5-FU-related toxicity occur during capecitabine activation, 5-FU activation, drug action, and/or drug catabolism.

Capecitabine Activation. One period where differences in enzyme levels or activity could lead to drug toxicity is during the activation of the oral prodrug capecitabine into 5-FU (Figure 1.1a). Two main causes of toxicity could exist during this period: complete activation of the drug prior to arrival at the colon and leakage of partially activated drug from the colonic tumour. Taken orally, capecitabine is absorbed by the gut as an inactive prodrug and is first processed to 5'dFCR by the enzymes *CES1* and *CES2*, almost exclusively in liver cells (144). 5'dFCR is then

metabolised to 5'dFUR via *CDA*, which is present mostly within liver cells but also in colon tumour cells (144). In the 5'dFCR or 5'dFUR forms, 5-FU metabolites are thus partially activated and hence require fewer enzymatic steps before becoming cytotoxic. As 5'dFCR and 5'dFUR are formed largely within the liver, it is possible that some toxicity results from conversion into 5-FU and its subsequent active forms (FdUMP, FdUTP and FUTP) prior to the drug arriving at the tumour site. Similarly, partially activated 5-FU that leaks from the tumour could become fully activated. However, these risks are mitigated by the fact that *UPP1* (necessary for conversion of 5'dFUR to 5-FU) expression is two- to three-fold higher in colon tumour tissue than in normal colon tissue, although this enzyme is also expressed in liver cells (126).

5-FU Activation. A second period that affects drug toxicity is during the activation of 5-FU (either provided directly by injection or as a capecitabine catabolite) to its active forms FdUMP, FdUTP and FUTP (Figure 1.1a/b). In the primary activation pathway, raw 5-FU and capecitabine metabolites that enter colon tumour cells are quickly metabolized to FdUMP, mostly via the *TYMP* pathway. Formation of FdUMP is concentrated within colon tumour cells because *TYMP* is more active in tumour cells than in adjacent normal tissue (144). Expression of *TK1* is similarly elevated in tumour cells (211). Various transporter proteins, including *ABCC4*, *ABCC5* and *ABCB1*, may also play a role in 5-FU toxicity by affecting levels of the drug within both the tumour and non-target tissues, such as via leakage. Additionally, 5-FU can be activated both within the tumour and in non-tumour tissues via the alternate activation pathways, primarily via *UMPS* (sometimes called *OPRT*) (127).

Drug Action. A third period affecting drug toxicity is during the mechanism of action of activated 5-FU (Figure 1.1a/b). 5-FU is not itself cytotoxic, but rather it is converted by enzymes into three active forms with distinct cytotoxic mechanisms (82; 127; 211):

1. FdUMP: binds to and inhibits the TYMS enzyme in place of binding by endogenous dUMP and thus blocks production of the DNA monomer precursor thymidine;

2. FdUTP: (mis)incorporates directly into DNA and overwhelms the excision repair mechanisms;

3. FUTP: (mis)incorporates directly into RNA and disrupts both pre- and post-translational modification.

Of note, the FdUMP/TYMS complex is stabilised by the 5,10-methylenetetrahydrofolate compound, a product of folate that is then reduced by the enzyme methylenetetrahydrofolate reductase (MTHFR) to 5-methyltetrahydrofolate (3; 150), meaning that elevated MTHFR levels might reduce the effectiveness of 5-FU.

5-FU Catabolism. Not all 5-FU becomes activated, but rather approximately 80% of injected 5-FU is catabolised prior to activation (84). The first enzyme in this catabolic pathway is DPYD. Thus, mechanism strongly suggests that mutations that affect the levels of DPYD activity should have a dramatic effect on drug toxicity. It is less clear whether mutations in enzymes downstream of DPYD would have a similar impact.

In summary, we see that while certain catabolic enzymes such as DPYD should have a clear mechanistic impact on 5-FU toxicity, the impact on toxicity is less clear for most enzymes, even while their roles in 5-FU activity and action are well established. This ambiguity is particularly so for *TYMS* and *MTHFR*. Mechanistically, we would expect low levels of these two enzymes to increase 5-FU-induced cytotoxicity (ie, drug response) and there is supporting evidence of this for *TYMS* (3; 160). However, it remains unclear whether increased or decreased germline levels of these enzymes should result in reduced levels of toxicity, and, importantly, whether the optimal enzymatic activity levels for toxicity align or conflict with those for efficacy.

For example, do low germline levels of drug target both increase efficacy by limiting DNA synthesis in tumour cells and increase toxicity by making non-tumour cells less resilient to drug exposure? This is hard to answer, as many enzymes are expressed both at the tumour site (allowing for efficacy with minimal toxicity) and in normal tissue (limiting efficacy and causing toxicity). Complicating the situation, while early stage tumours might largely reflect the germline levels of enzyme activity, later stage tumours might not, particularly following treatment with 5-FU, given resistance pathways in which 5-FU treatment can induce *TYMS* expression (152). Unfortunately, as a prior to genetic research for the 5-FU drug target enzymes, a mechanistic understanding does not allow us to predict directionality of DNA coding and regulatory alterations.

1.2.5 Rationale for Predictive Biomarkers of 5-FU Toxicity

Based on the previous three sections, there is a strong rationale for investigating whether germline genetic markers exist for 5-FU toxicity, namely:

1. Treatment benefit from 5-FU is modest, meaning that the risk/benefit balance of patient suffering exists narrowly.

2. Prevention of 5-FU toxicity would reduce patient suffering, given that one third of patients experience severe toxicity events, some of which result in death. Many of these patients have only months to live, particularly for mCRC. Others are being treated unnecessarily.

3. Prevention of 5-FU toxicity would further reduce healthcare costs, particularly those related to toxicity-mitigating interventions and patient travel to/from hospital.

4. Prevention of severe 5-FU toxicity cannot be accomplished merely by modifying the treatment of patients following early low-grade toxicity events, given

that toxicity occurs in "clusters." Prediction is necessary prior to first cycle of treatment.

5. Inter-patient differences in toxicity is a widely recognized phenomenon, and known factors such as age, gender and regional differences do not fully explain this variation.

6. There is a strong biological basis for how germline genetic polymorphisms might impact 5-FU pharmacokinetics and pharmacodynamics, particularly for enzymes involved in the drug's activation, action and catabolism, which would explain across-the-board 5-FU-related toxicity (ie, all phenotypes).

7. Some patients experience high levels of a particular toxicity phenotype, suggesting that polymorphisms might exist that impact a specific mechanism of toxicity (eg, the patient's predisposition for diarrhoea), possibly in genes for enzymes not yet known to impact 5-FU.

8. Various panels of genetic biomarkers could be clinically useful. For example, a panel that predicts dose-limiting levels of toxicity might alter treatment for 30% of patients, who would possibly be classified as "slow metabolisers" and started on lower doses of therapy. Or, a panel for extreme 5-FU-related toxicity could prevent the 0.5-3% of patients succumbing to 5-FU-related death, or a somewhat higher proportion who experience severely life threatening toxicities such as grade 4 neutropaenia.

An important counter-argument to the identification of biomarkers that predict 5-FU toxicity is the hypothesis that toxic drug levels benefit patient survival (ie, outcome), as has been observed in haematological malignancies (167). It is important, then, to identify both whether this is true generally for CRC, and, in the case that evidence supports this hypothesis, whether severe (ie, dose modifying) levels of

toxicity are necessary to achieve this benefit. Firstly, the evidence is generally clear that higher rates of 5-FU toxicity do not *reduce* patient survival, though statistical power can be limited. Returning to Stein et al, investigating bolus and infusional 5-FU patients, the authors found that stratifying their survival analysis by severe toxicity did not impact survival (186). A similar result was found by Cassidy et al for capecitabine; using Cox regression in which the onset of dose modification was a covariate and disease progression or death was the outcome variable, the authors found no increase in the hazard ratio upon dose modification of 25-50% in the capecitabine arm (HR=0.97, p=0.78) (33). On the contrary, these authors found that patients given bolus 5-FU/LV had increased risk of disease progression upon dose reduction, namely a 12% increased risk for a first dose reduction and a 36% increased risk for the second, though these results were not statistically significant. Further, in the long term survival follow-up to the X-ACT trial, Twelves et al found that handfoot syndrome in particular was associated with improved outcome from capecitabine treatment, namely that patients who experienced no HFS had similar survival to patients treated with bolus 5-FU/LV, whereas patients with grade 1 HFS or above experienced superior survival (73.8% vs 66.3% 5 year OS for stage III CRC; p-value and CIs not provided) (200). This supports the hypothesis that some patients may be fast 5-FU metabolisers and thus have reduced toxicity coinciding with reduced efficacy. Predicting this might allow for personalised dose. However, most toxicity phenotypes have little evidence of impacting outcome.

A second counter-argument to the identification of genetic markers of 5-FU toxicity is that many toxicity phenotypes can be safely managed. Considering a phase II study of FOLFOX for gastric cancer, among the more toxic 5-FU-based regimens, Louvet et al showed that, due to implementation of dose reduction, grade 3/4 neutropaenia and thrombocytopaenia were only experienced in a total (among all patients) of 8% and 1% of treatment cycles (128). None of the 11% of patients who

experienced febrile neutropaenia experienced recurrence after dose reduction. Of the 54 patients in this trial, one died from treatment-related toxicity. Thus while some risk of treatment-related death exists, for the most part, predictive markers of 5-FU toxicity could act to limit patient discomfort and cost, without minimizing efficacy, through use of predictive tests to identify the appropriate dose prior to treatment.

1.2.6 Genetic Biomarkers of 5-FU Toxicity

Investigation into genetic biomarkers of 5-FU toxicity is not new: a decade of publications exists. A summary of what has been reported in individual studies is provided here, while results for many of these studies are re-visited and meta-analysed in Chapter 3.

Historically, the first studies to report on the genetic causes of 5-FU toxicity linked severe cases (often deaths) to polymorphisms in *DPYD*, particularly the *2A variant, which causes deletion of exon 14 from mRNA (206). Following this, some 30 studies have retrospectively or prospectively used association testing to determine whether particular variants affect the incidence of 5-FU toxicity, mostly investigating known polymorphisms in candidate genes. Table 1.5 shows the 16 polymorphisms either for which there have been previously suggestive ($p<0.1$) reports or that are included in commercial diagnostic kits.

Association Testing

Certain biomarkers of 5-FU toxicity have been reported with modest confidence, such as *DPYD* *2A (173), *MTHFR* 1298 (2) and *TYMS* 5'VNTR (116). However, the existing published data on the genetics of 5-FU toxicity have been limited by three inconsistencies, namely in: the toxicity outcomes being reported, the pooling of patients on distinct 5-FU regimens and the formation of haplotypes (or suggested interactions) among polymorphisms. These inconsistencies make it difficult to deter-

	Gene symbol(s)	Polymorphism(s) (commonly used identifiers)	Most significant previous association
Activation	CES2	rs8192924 (6046G>A; R270H)	p=0.09 with diarrhoea for CAP (Ribelles)
	CDA (CDD)	rs532545 (-451C>T)	p=0.039 with HFS for CAP (Caronia)
		rs3215400 (943insC)	p=0.028 with HFS for CAP (Caronia)
Target	TYMS (TS)	rs45445694 (G), rs34743033 (C) (5'UTR TSER 2R/3R; 5'VNTR)	p=0.019 with global for BOL or INF (Schwab) included in commercial kits
		rs2853542 (5'UTR TSER 3R G/C SNP)	p=0.09 with global for BOL (Gusella)
		rs16430 (3'UTR 1494indel6b)	p=0.012 in haplotype with global for BOL or INF (Lecomte) p=0.10 alone with global for BOL (Gusella) included in commercial kits
	MTHFR	rs1801133 (677C>T; A222V)	p=0.08 alone with global (Gusella) included in commercial kits
		rs1801131 (1298A>C; E429A)	p=0.002 in haplotype with TYMS 3'UTR with GI toxicity for BOL (Afzal) included in commercial kits
Catabolism	DPYD (DPD)	-1590 T>C	no formal statistical testing included in commercial kits
		rs1801265 (DPYD *9A; 85T>C; C29R)	p=0.05 with vomiting for INF+OX (McLeod) included in commercial kits
		rs1801158 (DPYD *4A; 1601G>A; S534N)	reported non-significant with global for BOL or INF (Schwab) included in commercial kits
		rs1801159 (DPYD *5; 1627A>G; I543V)	p=0.04 with diarrhoea for INF+OX (McLeod) included in commercial kits
		rs55886062 (DPYD *13; 1679T>G; I560S)	no formal statistical testing included in commercial kits
		rs3918290 (DPYD *2A; IVS14+1G>A; 1905+1G>A)	p=0.01 with global for BOL or INF (Schwab) included in commercial kits
		rs1801160 (DPYD *6; 2194G>A; V732I)	reported non-significant with global for BOL or INF (Schwab) reported non-significant with global for INF + OX or IR (McLeod) included in commercial kits
		rs67376798 (2846T>A; D949V)	reported significant with global for BOL or INF (Schwab) included in commercial kits

Table 1.5: Since 1999, dozens of polymorphisms have been tested for an association with 5-FU toxicity. Of these, 16 might be considered true associations upon reviewing the literature, as they have either shown a suggestive p-value or have been included in commercial diagnostic kits that aim to predict 5-FU toxicity. SNPs are listed by gene and for each, the most significant association found in the literature is reported. Strikingly, there is little consistency in both the toxicity phenotypes for which associations are reported and the particular variants reported within a given gene. Afzal=(2), Caronia=(31), Gusella=(79), Lecomte=(116), McLeod=(140), Schwab=(173).

mine whether the effects seen in one study truly match those reported in either a positive or negative validation study that follows, and thus whether a biomarker is both predictive in a tested regimen and truly generalisable to the population at large.

Toxicity Outcomes. Outcome variables have shown inconsistency in both the grouping of toxicity phenotypes and the grouping of toxicity event grades. Some authors, for example, report that polymorphisms have an effect on individual toxicity phenotypes (eg, neutropaenia or diarrhoea), while others group these toxicity phenotypes into measures such as "haematologic," "gastrointestinal," "non-haematologic," and "non-gastrointestinal." Within these categories, the toxicity scores can be further grouped, sometimes reporting grade 0 vs 1 vs 2 vs 3 vs 4 toxicity, but other times reporting grade 012 vs 34 toxicity. Of course, as described in Section 1.2.2, there are clinical reasons why lower grade levels of certain toxicities are as important to predict as higher grade levels of other toxicities. However, even within a single toxicity phenotype, reporting of grades differs. This leads to confusing interpretation of the field. For example, while Sharma et al report that the *MTHFR* 677 T-allele and 1298 A-allele are predictive of any *grade 2/3* capecitabine toxicity (n=54, $p<0.05$ and $p<0.01$) (175), Largillier et al report that neither of these variants are predictive of any *grade 3/4* capecitabine toxicity (n=84, p=0.54 and p=1.00) (115). This is not to mention that Sharma et al use a recessive model and mCRC patients, while Largillier reports an allelic model and uses breast cancer patients (though drug doses are similar), and that the list of included toxicity phenotypes differ slightly. Neither approach is necessarily right, but the inconsistencies make interpretation quite difficult.

Patient Pooling. Differences in toxicity outcomes among small studies could be mitigated by looking to the large studies, however, the patients in these large studies are often a heterogeneous group in regards to the treatments that they receive. Naturally, the data available for most biomarker studies come from translational studies of clinical trials, and the goal of most such trials is to compare two or more treat-

ment options. Unfortunately, however, while treatment effect is reported separately (and compared) across 5-FU regimens, biomarker results are sometimes pooled into a single analysis of "any patient receiving a 5-FU-based regimen," and sensitivity analyses stratified by regimen are limited. Take, for example, the *TYMS* 5'VNTR polymorphism and its effect on global grade 3/4 toxicity. In Schwab et al (173), this polymorphism is reported on for 521 patients receiving one of seven 5-FU-based monotherapy regimens ranging from high-dose infusion to weekly short infusion to continuous infusion, whereas in Braun et al (25), the variant is reported on for 688 patients receiving infusional 5-FU, 270 receiving FOLFIRI and 280 receiving FOLFOX. Thankfully, Braun et al stratify their analysis by regimen; however, even here the results across studies remain difficult to interpret: do we consider the Braun et al result (p=0.40 for infusional 5-FU; no OR provided) to be negative validation of the Schwab finding (OR=0.52, 95% CI 0.34-0.81, p<0.01)? Or do we instead believe that a certain subpopulation of patients treated within Schwab et al (but not receiving the same infusional 5-FU monotherapy regimen as Braun patients) are driving away an effect that would otherwise be shown among infusional 5-FU patients? The pooling of patients can be further observed by running down the righthand column of Table 1.5. Again, each author provided the results that made the most sense given his data, but the answer at large remains unclear.

Haplotypes, Gene-Gene Interaction and Model. A third source of inconsistency is in the haplotype formation or joint analysis among previously reported polymorphisms and the particular model (ie, allelic, recessive, dominant) used to analyse these combined polymorphisms. Some haplotypes have a biological rationale, and thus make sense to consider testing. The clear example for this is the G/C SNP present on the third repeat of the *TYMS* 5'VNTR polymorphism, which previous functional work has suggested affects the total number of USF1/USF2 binding sites across both the repeat and G/C alleles (131). Other polymorphisms have been com-

bined with less prior justification. For example, prior work has shown that there is a gene-gene interaction between *MTHFR* and *TYMS* affecting folate metabolism, using the *MTHFR* 677 and *TYMS* 5'VNTR polymorphisms (105). Subsequently, an interaction between these genes was reported by Afzal et al, however, only between different polymorphisms in these genes – namely that the *MTHFR* 1298 A-allele and *TYMS* 3'UTR del-allele increase one's risk of specifically gastrointestinal 5-FU toxicity (exploration cohort: OR=2.40, p=0.003; validation cohort: OR=1.81, p=0.007) (2). Given that 11 polymorphisms were assessed, such reports raise concern of post-hoc reporting decisions. Similar concerns exist when certain polymorphisms within a gene are reported without justification as significant under an allelic model and others under a recessive model, as occurs for *MTHFR*.

In addition to these main inconsistencies, it is also important to consider whether the generalisability of results is further limited by differences in disease stage, ethnicity, DNA collection source (ie, blood vs tumour) and genotyping method, toxicity scoring system, prospective vs retrospective patient recruitment, and study sample size. Together, these inconsistencies suggest that post hoc reporting decisions may bias our understanding of the field. Thus, by simply reviewing the literature, it is hard to be truly convinced that any particular polymorphism will be useful in predicting 5-FU toxicity for any particular individual.

Functional Evidence

For some variants – in particular those that are rare – it will be easier to identify a predictive capacity for toxicity by performing functional studies than it will be via association testing. For many of the 16 polymorphisms highlighted in Table 1.5, some degree of functional evidence exists. For other variants, there exists functional evidence but limited association testing. Table 1.6 summarises the existing functional evidence for variants proposed to be predictive of 5-FU toxicity. Note that this table

excludes functional evidence for most rare *DPYD* variants thought to be causal of DPYD Deficiency Syndrome (OMIM 274270), discussed further in Chapter 3.

Strong evidence for a functional role in 5-FU toxicity would take the form of an assay in which both mutant and wildtype enzyme is produced *in vitro*, using DNA that is identical except for the variant in question, and tested to see whether differences exist in the metabolism of 5-FU. This evidence is particularly strong when the variant is rare and has been observed as a homozygote or compound heterozygote (with a known causal variant) in patients with DPYD Deficiency Syndrome, though it is still not perfect because it assumes that 5-FU catabolism correlates with toxicity levels. Only in 2013, the first such controlled *in vitro* assay was applied to the study of 5-FU toxicity (149). In this experiment, *DPYD* was first cloned into cells with no endogenous *DPYD* gene, then constructs were formed for six variants of interest using site-directed mutagenesis. Finally, enzyme activity was determined by adding 5-FU to cells and quantifying its conversion to a catabolite, using flow scintillation analysis.

Suggestive functional evidence could take a few forms, such as identifying modest differences when using the assay just described, or by identifying constituent levels of DPYD activity in blood of patients with various genotypes. This latter assay is a weaker option because it allows for extra variables, given the *in vivo* design; for example, additional functional polymorphisms could act as confounders, as could all of the clinical factors mentioned in Section 1.2.3. Such a method was used by Seck et al (174), who investigated 23 *DPYD* variants using previously described methods (107; 129) in which HPLC was used to measure the amount of a DPYD catabolite present after adding isolated DPYD from blood to a 5-FU solution. In Discussion, these authors point out that the tested patients exhibited minimal variation in enzyme activity, which they suggest is because Caucasians have minimal DPYD deficiency; however, this may also point to a limited sensitivity of the assay, due to the con-

Table 1.6: Functional evidence for 5-FU toxicity variants. Functional evidence exists for 17 putative 5-FU toxicity variants, of which 8 are not in *DPYD*. Evidence is of varying impact on our ability to conclusively view a variant as predictive of 5-FU toxicity. DPYD Deficiency Syndrome (OMIM 274270) is a rare and heterogenous phenotype caused by functional homozygous or compound heterozygous mutation of *DPYD*, sometimes leading to mental retardation or convulsive disorder. Caronia 2011=(31), Cecchin 2005=(34), Frosst 1995=(67), Ghosh 2011=(69), Mandola 2003=(131), Mandola 2004=(132), Offer 2013=(149), Seck 2005=(174), VanKuilenburg 2000=(204), VanKuilenburg 2002=(203), Vreken 1997=(207), Weisberg 1998=(209), Wu 2004=(215), Zhang 2011=(222).

Polymorphism (Annotation)	**Functional Evidence for 5-FU Toxicity**
***CES2* 823C>G** (promoter)	**Inconclusive.** Haplotypes with the homozygous 823 G-allele have lower mRNA levels in human blood but no corresponding change to substrate activation **(Wu, 2004)**; other work has shown an association between lower CES2 mRNA levels and a decrease to substrate activation, though 823 polymorphism and capecitabine/5-FU not studied **(Cecchin, 2005)**.
***CDA* -451C>T** (promoter)	**None. Caronia et al 2011** found no correlation between genotype of this SNP and CDA mRNA expression in a panel of 89 lymphoblastiod cell lines (p=0.671). Strong linkage disequilibrium (D-prime=0.92) with CDA943insC, which may drive effect on expression.
***CDA* 943insC** (5'UTR)	**Suggestive. Caronia et al 2011** performed *in silico* analysis and suggest that the inserted C is required for E2F binding. In a panel of 89 lymphoblastoid cell line, homozygotes for the inserted C allele had significantly lower mRNA levels compared to wild type homozygotes and heterozygotes (p=0.004).
***TYMS* 5'VNTR 2R/3R** (promoter)	**Minimal. Zhang et al 2011** found that pancreatic cell lines homozygous for the 3R allele (n=2) expressed higher levels of TYMS protein (western blot) compared to cell lines with the 2R/3R or 2R/2R genotypes (n=5). The 3R homozygous cell lines also had lower growth inhibition rates compared to 2R/3R and 2R/2R cell lines when treated with 5-FU. **Ghosh et al 2011** found that while the 3R led to slightly lower protein levels in a firefly construct, no difference was observed using a human construct, using human CRC cell preparations (n=7).
***TYMS* 5'VNTR 3R G/C SNP** (promoter)	**Minimal. Zhang et al 2011** found that 4 pancreatic cell lines of either 3RC/3RC (n=1) or 2R/3RC (n=3) had lower protein levels of TYMS compared to 1 3RG/3RG cell line. The difference between protein levels in the 3RC/3RC and 2R/3RC was significant. **Mandola et al 2003** showed that phosphorylated USF-1 binds to the 28 repeat sequence only when the G allele is present.
***TYMS* 3'UTR 1494indel6bp** (promoter)	**Suggestive. Mandola et al 2004** measured TYMS mRNA levels in 43 tumours. mRNA levels were significantly lower in tumours with the del/del genotype compared to those with the 6bp/6bp genotype (p=0.007). Luciferase reporter constructs containing the 6bp ins allele and the del allele had reduced luciferase acitivity, by 35% and 70% respectively.
***MTHFR* 677C>T** (non-synonomous (A>V))	**Suggestive. Frosst et al (1995)** found that MTHFR activity was statistically significantly lower in lymphocytes from patients carrying one or more copies of the T allele (N=40). This finding was validated by **Weisberg et al 1998** in 174 Canadian subjects.
***MTHFR* 1298A>C** (non-synonomous (E>A))	**Suggestive. Weisberg et al 1998** showed that carriers of the 1298 C-allele had lower MTHFR activity but that the effect size was not as large as that of 677T.

Polymorphism (Annotation)	**Functional Evidence for 5-FU Toxicity**
***DPYD* 85T>C** (non-synonomous (C>R))	**Inconclusive.** Minimal functional effect seen among 56 non-WT blood donors **(Seck, 2005)**, but present in patients with partial DPYD Deficiency Syndrome **(van Kuilenburg et al 2000)**. Shown to reduce enzyme activity in E. coli **(Vreken, 1997)**, however, not using simple site directed mutagenesis method, and shown to increase DPYD activity by 13% in mammalian cell lines by **Offer et al 2013** (p=0.0013).
***DPYD* 496A>G** (non-synonomous (M>V))	**Modest**. No functional effect seen among 24 non-WT blood donors **(Seck, 2005)**, but present in patients with partial DPYD Deficiency Syndrome **(van Kuilenburg et al 2000)**.
***DPYD* 1236G>A** (synonomous (E>E))	**None.** No functional effect seen among 1 non-WT blood donor **(Seck, 2005)**.
***DPYD* 1601G>A** (non-synonomous (S>N))	**Inconclusive.** Decreased DPYD activity observed among 5 non-WT blood donors by **Seck et al 2005** (p<0.05). 36% increased DPYD activity observed by **Offer et al 2013** (p=3.4e-7).
***DPYD* 1627A>G** (non-synonomous (I>V))	**Weak.** Minimal functional effect seen among 40 non-WT blood donors **(Seck, 2005)**. Present in patients with partial DPYD Deficiency Syndrome **(van Kuilenburg et al 2000)**, though the variant is common (MAF=20%). No effect on DPYD activity observed by **Offer et al 2013**.
***DPYD* 1679T>G** (non-synonomous (I>S))	**Strong. Offer et al 2013** showed a 75% reduction in activity relative to WT (p=5.2e-9). Present as a compound heterozygote in a DPYD Deficiency Syndrome patient **(Van Kuilenburg, 2002).**
***DPYD* *2A IVS14+1G>A** (essential splice site, A-allele results in skipping of exon 14)	**Strong.** In a cohort of 60 cancer patients experiencing grade 3/4 toxicity after administration of 5-FU, 17 were heterozygous or homozygous carriers of the A-allele and 16 of 17 had low DPYD activity **(Van Kuilenburg, 2002)**. Present in patients with partial DPYD Deficiency Syndrome **(van Kuilenburg et al 2000)**. In mammalian cell lines deficient in DPYD, this variant led to the same expression levels as those of empty vectors **(Offer, 2013)**.
***DPYD* 2194G>A** (non-synonomous (V>I))	**Inconclusive.** No functional effect seen among 7 non-WT blood donors **(Seck, 2005).** Present in patients with partial DPYD Deficiency Syndrome **(van Kuilenburg et al 2000)**. No decrease to enzyme activity observed in mammalian cell lines **(Offer, 2013)**.
***DPYD* 2846T>A** (nonsynonomous D>V)	**Suggestive.** Carriers of the 2846 A-allele had lower DPYD activity **(Seck 2005)**; of note, the patients did not have lower *DPYD* mRNA levels, suggesting that the expression level of the gene is not the cause of the lower activity. Present in patients with partial DPYD Deficiency Syndrome **(van Kuilenburg et al 2000)**.

founders mentioned above. One improvement on this experiment would have been if the authors further correlated the enzyme activity levels with toxicity levels.

Weaker functional evidence might show a correlation between a polymorphism and a given cellular phenotype, but without necessarily demonstrating that this cellular phenotype has an effect on 5-FU toxicity. Moving beyond DPYD, for example, Mandola et al (131) showed that the G/C SNP in the *TYMS* 5'VNTR third repeat can abolish the ability of USF-1 (a transcription factor) to bind. It is not clear, however, that abolishing this site truly impacts *TYMS* expression, and furthermore whether

this impacts 5-FU toxicity, though it does suggest that an association study between genotype and toxicity might be worthwhile.

Of the 17 putative 5-FU toxicity variants with functional reports, only two have strong evidence of impacting levels of 5-FU toxicity, and even these had only modest functional evidence prior to 2013. Furthermore, this strong functional evidence is limited to *DPYD*, meaning that we still do not know whether varying levels of other candidate and non-candidate enzymes impact risk of 5-FU toxicity, much less whether polymorphisms in these variants have a functional effect. As discussed in the introduction to Chapter 4, the functional evidence that links *TYMS* genotype (particularly of the UTR polymorphisms), expression and toxicity remain controversial.

Commercial Kits for the Prediction of 5-FU-Related Toxicity

Importantly, despite this level of genetic and functional evidence, many polymorphisms have been included in commercial kits that predict 5-FU toxicity based on genetic biomarkers. These kits do not provide estimates of their sensitivity, specificity and negative and positive predictive power. Rather, instead of validating a set of potential markers on a cohort of patients, at least openly, they rely on recommendations from the literature, as is apparent from their technical specifications. These commercial kits range from those that merely provide genotyping of the *DPYD* *2A polymorphism, to those genotyping polymorphisms of multiple genes and performing exon sequencing, as highlighted in Figure 1.2.

In summary, key questions remain as to which polymorphisms act as true predictors of 5-FU toxicity, and for which patients. The field is not only faced with a vast number of variables to consider regarding patient characteristics, but also a range of outcome variables to consider testing. Thus, a key methods question is inherent to the further investigation of makers of 5-FU toxicity – namely how to balance the need for broad discovery-phase testing and the problems that arise out of multiple testing.

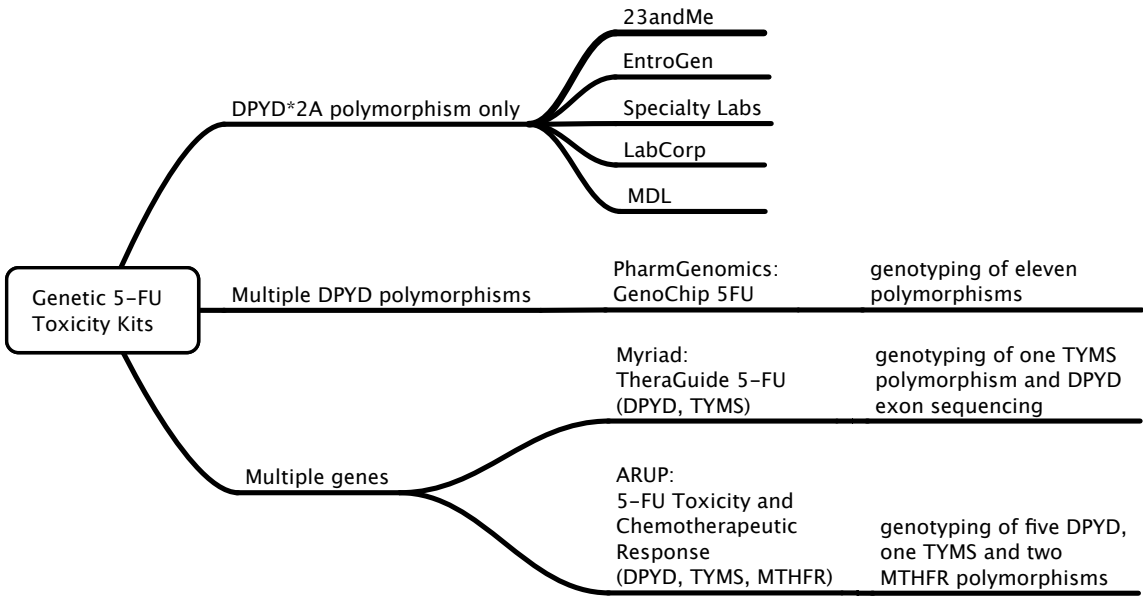


Figure 1.2: Various commercial kits exist that aim to predict 5-FU toxicity based on genetic markers. Generally, these kits rely on the literature as a source of polymorphisms to include, though many of these variants are not validated and the predictive power of these kits remains unknown.

1.2.7 Non-genetic Biomarkers of 5-FU Toxicity

Non-genetic biomarkers are a potentially useful means of predicting 5-FU toxicity, either alone or in combination with genetic markers. Primarily, these take the form of DPYD enzyme activity measurements, either directly or indirectly. Indirect measurement is more common, measuring, for example, the uracil levels in one's blood (20). On the one hand, measurement of actual enzyme activity levels avoids the confounders present when genotyping only a subset of one's genome; however, the methods are generally expensive, particularly when applied to multiple genes (whereas genotyping costs decrease rapidly with scale, allowing for high-throughput screening).

1.3 CRC Prognosis

The spectrum of care and treatment provided for an individual patient is based on myriad factors, among which prognosis (ie, survival post diagnosis) is one, alongside

likelihood of efficacious response and propensity for dose limiting toxicity. Historically, these decisions are made based on clinical factors, such as disease stage and performance status. However, the personalization of CRC treatment based on an individual's germline DNA may improve treatment in terms of efficacy, safety and patient comfort.

1.3.1 Clinical Factors Affecting CRC Prognosis

Clinical factors drive much of the inter-patient difference in CRC survival.

Disease stage is the most significant. Beyond the thousands of pre-cancerous polyps that are removed annually, approximately 10% of patients present with superficial lesions (stage 0), 40% present with localized disease (stage 1 and early stage 2), 35% present with advanced regional disease (late stage 2 and stage 3), and 15% present with distant/metastatic disease (stage 4). As shown in Table 1.7, survival differs dramatically by stage, but the histopathological features used to determine stage are imperfect alone. Among these cases, patients with advanced and metastatic disease require chemotherapy. Possibly, genetic biomarkers may be able to further explain disease progression, as well as the inter-individual differences in survival within each stage.

CRC is also a disease that differs with age. According to the English Office for National Statistics, 5-year survival decreases with each decade of life, both for males and females – in one's 40s, CRC 5-year survival is approximately 60%, falling to just over 40% in one's 80s. This age effect exists even when adjusting for the increased background for age-induced mortality. Although CRC incidence differs by gender, sex has little effect on survival rates. Again according to the English Office for National Statistics, females have slightly higher relative survival in each age bracket, though not significantly so. Likewise, other known CRC risk factors have demonstrated little impact on CRC survival. Based on these clinical factors, previous analyses of the

Table 1.7: CRC 5-year overall survival rates are shown by AJCC stage according to TNM (tumour/node/metastasis), using data from the NCI SEER database of more than 28,000 patients diagnosed between 1998 and 2000 (`http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-survival-rates`). Of note, the continuum of histopathological features that determine stage result in certain "later stage" cancers having a better prognosis than some "earlier stage" cancers.

Stage	Survival
I	74%
IIA	67%
IIB	59%
IIC	37%
IIIA	73%
IIIB	46%
IIIC	28%
IV	6%

genetic markers of CRC survival have adjusted for three clinical factors: age, stage and gender (114; 192).

Somatic (ie, tumour) DNA changes help further sub-classify some patients. Genomic instability provides one source. Chromosomal instability (CIN) has shown a strong effect on prognosis by meta-analysis, with CIN+ (aneuploid/polyploid) patients experiencing far worse overall survival (HR=1.45, p<0.001, n=10,126), although the biologic rationale is not certain (208). Similarly, microsatellite instability (MSI) has been shown to improve patient survival by meta-analysis (HR=0.65, 95% CI 0.59-0.71, n=7,642), meaning that patients with functional mismatch repair proteins suffer from poor prognosis (156). CIN and MSI are negatively correlated, with only about 2% of tumours displaying both instabilities (56). Finally, loss of chromosome arm 18q has been implicated in poor CRC prognosis by meta-analysis (HR=2.00, 95% CI 1.49-2.69, n=2,189) (155), though not all studies show an effect.

Specific gene mutations are another source. Mutations in *KRAS*, particularly codon 12 glycine to valine, have been associated with poor prognosis (OS, HR 1.29,

p=0.008) (8), though not in all studies. It is possible that *BRAF* mutations similarly affect this pathway (164). Like for the germline markers described in the following section, though, these mutations may show more of a treatment-interaction effect than an inherent underlying cause of inter-individual differences in disease progression (158).

While non-germline genetic factors drive much of the inter-patient differences in survival, they do not explain all of the variation, which is important toward guiding the appropriate treatment decisions.

1.3.2 Germline Genetic Biomarkers of CRC Prognosis

The study of germline genetic biomarkers of CRC survival is less developed than the study of 5-FU toxicity, including for germline markers. In a sense, this is a surprise – the two investigations have much in common, namely that their study is inherently tied to clinical trial datasets. However, survival analysis requires 3- or 5-year follow-up data, meaning results are delayed or sometimes not collected. Further, while events (ie, deaths) are frequent, the different interventions used across trials means that datasets are not easily combined in order to increase statistical power, and thus most results cannot be interpreted as affecting general CRC aetiology.

The most common CRC prognosis analyses have been those of drug response, testing whether SNPs within candidate genes for drug efficacy (eg, drug target proteins) also might impact patient survival. Some have been combination-regimen analyses, such as for cetuximab (223) or bevacizumab (65). From a clinical perspective, such studies are only relevant to a subset of CRC patients. From a scientific perspective, the specific characteristics of included patients makes these studies hard to replicate and does not provide a general understanding of CRC aetiology. Studies of prognosis among patients given 5-fluorouracil monotherapy regimens are more broadly applicable. However, like for 5-FU toxicity, these prognosis studies have shown conflicting

results, possibly complicated by differing effect in the adjuvant and metastatic setting, and have been limited to a focus on candidate genes, mostly of 5-FU toxicity enzymes such as *TMYS* (57; 136). Of note, the *TYMS* polymorphisms were more significant when analysing tumour DNA than when analysing DNA from surrounding normal tissue, further limiting our understanding of germline markers of CRC survival. More broadly applicable is the recent finding that aspirin use improves CRC survival among the subset (approximately 15%) of patients with *PIK3CA* mutations (122).

Additional studies have investigated genes beyond those related to drug effects, though these have been limited for other reasons. Iglesias et al showed that a variant within the *COX2* gene improves survival (p=0.03), but used only 284 patients (number of events not provided) and investigated only five polymorphisms (95). Koessler et al identified multiple polymorphisms within mismatch repair genes to be associated with overall and disease-specific survival (114). This study was large (n>2000), studied 68 tagging SNPs and provided good significance (p=0.001 for some SNPs). However, the study only focused on a limited number of genes, selected for their previously studied role in CRC predisposition.

Others have searched more broadly within CRC susceptibility loci, though little has been found. Xing et al found that rs10795668 (10p14) was associated with reduced risk of recurrence, but only for patients treated with chemotherapy (216); this is a relevant finding, but again limits our knowledge to drug effect. Cicek et al investigated five 8q24 CRC susceptibility loci, but found no associations (38). Furthering this report, Tenesa et al showed that none of the 10 GWAS-identified CRC susceptibility loci as of 2010 (90) showed an effect on CRC survival after diagnosis, using over 2,800 patients (947 deaths, 811 from CRC) (192). In some contrast to this, Passarelli et al showed that certain alleles of two SMAD7 (18q21) SNPs (rs4939827 and rs4464148) affected survival, though only among recent (ie, close to diagnosis) pre-diagnosis

NSAID users; further, sample size was limited to 727 patients (151). The only positive study across patient groups has been Phipps et al, who studied all 16 GWAS-identified CRC susceptibility loci identified as of 2011 (196) and showed that the minor allele of rs4939827, which is associated with lower risk of CRC incidence, reduced overall CRC survival (HR=1.16, p=0.002) among 2,611 patients in five cohorts (153). Despite meta-analysis, it is unclear whether this result is impacted by reporting bias, given that the collection of data was not performed systematically – for example, it failed to include the previously published cohort from Tenesa et al; the direction of effect in this study, however, did align. Similar analyses in breast cancer have also identified associations with survival among predisposition loci (16).

In summary, while the study of germline markers of CRC prognosis has expanded to include a broad group of CRC patients, most identified markers are specific to particular treatments. Possibly, genetic inter-individual differences are driven mostly by such interactions. However, the candidate-based scope of analyses thus far used means that our understanding is limited to drug effects and known CRC predisposition variants.

1.4 Approaches to Genetic Biomarker Analysis

Based on the nature of the dataset(s) available, the proper course of analysis and required external resources can differ greatly. This section provides background for the analysis approaches used throughout this thesis, while exact methods used are described in Methods.

1.4.1 Meta-analysis

Given the large body of literature regarding genetic biomarkers of 5-FU toxicity, one useful investigation might be to systematically review and meta-analyse the existing

studies. Such an approach is common in the medical sciences, but generally is applied to testing the effect of a particular intervention, not the effect of a particular allele. However, the methods of meta-analysis can be applied to either study goal (221), and the effect of various toxicity alleles are meta-analysed in Chapter 3.

The purpose of meta-analysis is to integrate data from multiple sources to answer a study question (49), often that requires more data than any individual study can provide. To accomplish this, the method applies five steps: forming a study question, identifying relevant studies, evaluating found studies, extracting results, and analysing extracted data (18). These steps follow rigid protocols so as to avoid introducing bias; the Cochrane Collaboration (`http://www.cochrane.org/`) provides guidelines for methods, while PRISMA (`http://www.prisma-statement.org/`) provides guidelines for reporting on meta-analysis results. The principles of meta-analysis can be found at `www.cochrane-net.org/openlearning/HTML/mod0.htm` in the 21-part "Cochrane Collaboration open learning material," used as the source for much of this section.

While the design and data aggregation steps are often the most time intensive, the purpose for completing a meta-analysis flows from its mathematical approach. This is best explained by comparing meta-analysis to two alternatives, all three of which aim to determine the effect of a certain exposure (eg, a drug or genotype) on case/control status (eg, sick/healthy). The first is a review article, in which previous publications are collected and summarised individually. The key difference is that meta-analysis provides a single effect size across all collected data, while a review merely reports on and interprets the individual effect sizes observed in each study. The second is a pooled analysis, in which data from previous studies are combined into a single dataset (ie, one large study is formed from many) and then analysed. Meta-analysis is similar in that the data from the separate studies are combined, however, meta-analysis calculates the single effect size from individual studies and

then pools these estimates, instead of pooling the data from individual studies and calculating the effect size thereafter. To do so, meta-analysis weights the effect size observed in each study by either the study's sample size or, in other methods, inverse variance, which accounts for both sample size and event rate.

This last point is critical, because in certain situations the result obtained from pooling then analysing fundamentally differs from the result obtained from pooling analyses. This difference is known as Simpson's Paradox, for which the seminal paper described how mathematically a dataset that shows a positive treatment effect on both males and females individually can show a negative treatment effect when pooling males and females into one dataset (178). Table 1.8 provides theoretical data to elucidate this point, as it applies to the study of genetic biomarkers of toxicity. In (A), we see that when two identical studies are pooled into a single study, the effect estimate (OR) is the same when either pooling the data then analysing, or analysing the individual effect estimates and then meta-analysing the odds ratios. Similarly, when the data from two studies of similar size with inverse effect estimates and the same variance are pooled (B), the effect estimate is the same as that from meta-analysis. Why, then, use meta-analysis? In less contrived datasets (C), we see that the effect estimate by pooling and meta-analysis differs slightly; however, who is to argue which is more correct? The answer lies in the fact that data pooling fails to provide a rational result when included studies of different sizes have odds in different directions (ie, above/below 1), but odds ratios in the same direction. This, shown in (D), is Simpson's Paradox. Clearly, though, there should be careful consideration as to whether Study 1 and Study 2 in (D) should even be evaluated together, given the differences in allele frequency and toxicity rates needed to make the Paradox play out. For reasons such as this, there is much debate as to whether meta-analysis is the appropriate method for answering clinical questions (12; 35; 118; 193).

Table 1.8: For many sets of study data, data pooling and meta-analysis provide a similar, if not identical, estimate of the overall effect estimate. However, meta-analysis provides a better estimate when the data underlying odds ratios of various studies differ in terms of their exposed/unexposed and case/control ratios. Part (D) expanded upon and adapted partially from the "Cochrane Collaboration open learning materials." HT=high toxicity patient, LT=low toxicity patient, MAF=minor allele frequency, OR=odds ratio.

(A) Same Effect Size

STUDY 1	HT	LT	Odds	OR	MAF	% HT
Allele A	16	19	1.19	**0.55**	0.65	0.41
Allele B	6	13	2.17			
STUDY 2	HT	LT	Odds	OR	MAF	% HT
Allele A	16	19	1.19	**0.55**	0.65	0.41
Allele B	6	13	2.17			
POOLED	HT	LT	Odds	OR	MAF	% HT
Allele A	32	38	1.19	**0.55**	0.65	0.41
Allele B	12	26	2.17			
META-ANALYSED				OR		
				0.55		

(B) Inverse Effect Sizes

STUDY 1	HT	LT	Odds	OR	MAF	% HT
Allele A	16	19	1.19	**0.55**	0.65	0.41
Allele B	6	13	2.17			
STUDY 2	HT	LT	Odds	OR	MAF	% HT
Allele A	19	16	0.84	**1.82**	0.65	0.59
Allele B	13	6	0.46			
POOLED	HT	LT	Odds	OR	MAF	% HT
Allele A	35	35	1.00	**1.00**	0.65	0.50
Allele B	19	19	1.00			
META-ANALYSED				OR		
				1.00		

(C) Realistic Data

STUDY 1	HT	LT	Odds	OR	MAF	% HT
Allele A	16	19	1.19	**0.55**	0.65	0.41
Allele B	6	13	2.17			
STUDY 2	HT	LT	Odds	OR	MAF	% HT
Allele A	36	39	1.08	**0.85**	0.56	0.46
Allele B	26	33	1.27			
POOLED	HT	LT	Odds	OR	MAF	% HT
Allele A	52	58	1.12	**0.78**	0.59	0.45
Allele B	32	46	1.44			
META-ANALYSED				OR		
				0.76		

(D) Simpson's Paradox

STUDY 1	HT	LT	Odds	OR	MAF	% HT
Allele A	16	19	1.19	**0.55**	0.65	0.41
Allele B	6	13	2.17			
STUDY 2	HT	LT	Odds	OR	MAF	% HT
Allele A	45	5	0.11	**0.69**	0.43	0.88
Allele B	56	9	0.16			
POOLED	HT	LT	Odds	OR	MAF	% HT
Allele A	61	24	0.39	**1.11**	0.50	0.73
Allele B	62	22	0.35			
META-ANALYSED				OR		
				0.62		

1.4.2 Candidate-Based Analyses

Candidate and pathway-based analyses test polymorphisms in and around genes with a suspected role in the aetiology of a disease (190). This method allows researchers to search for novel functional loci and tagging SNPs while reducing the multiple testing burden that can mask true signals among false positive of modest p-value. This ability

is particularly important for alleles of modest (eg, OR=1.5) effect size (44). As such, candidate and pathway-wide analyses have proven fruitful both in identifying new risk loci for disorders (13; 180; 196) and distancing certain pathways once considered causal (109).

This reduced testing burden comes at a cost, however, as it is particularly prone to publication bias toward false positives on a literature-wide basis. This has been demonstrated, for example, with major depressive disorder, in which Bosker et al identified 57 candidate genes from the literature with at least one significant association and validated the results with a study whose sample size was ten time larger than the mean study size; associations in only four of the genes replicated positively, and only did so minimally (23). This phenomenon has been recognized widely; across all areas of epidemiologic study, Ioannidis et al estimate that while the false positive to false negative finding ratio of GWAS is 1:100 or lower, that of candidate studies can exceed 100:1 (101). Some means of mitigating this risk would be to either focus a candidate-based analysis on genes containing SNPs that have previously reached genome-wide significance, or to correct for multiple testing across all SNPs in a broad set of candidate genes (ie, not trust reports investigating single candidate genes when multiple genes of similar interest exist).

Candidate-based analyses, however, are inherently limited to an hypothesis-driven approach, limiting their power in complex traits (6).

1.4.3 Genome-Wide Association Studies

The cheapening of microarray-based technology and computational power has allowed for the growth of non-hypothesis-driven approaches to genetic discovery, one of which is SNP GWAS. Five years after hitting its stride, approximately 1,500 GWAS are published and the method has identified hundreds of replicated disease-associated SNPs (189) (as maintained at `http://www.genome.gov/gwastudies/`). In addition

to technology, this explosion has been fueled by the expansion of open-access resources such as the 1000 Genomes Project (68) and Wellcome Trust Case Control Consortium (210), which have in turn led to improved *in silico* methodology such as imputation (134) and ENCODE (40).

The method is not without criticism, however, as it both allows for the identification of false positives and has failed to identify the extent of disease heritability once presumed to be within its capacity. In his 2012 review, Gibson summarises where this "missing heritability" may be hiding: "[as] a large number of small-effect common variants across the entire allele frequency spectrum (the infinitesimal model), a large number of large-effect rare variants (the rare allele model), or some combination of genotypic, environmental and epigenetic interactions (the broad sense heritability model)" (70). In response to the second of these criticisms, follow-on studies using next generation sequencing technology have been used in order to fine map regions of interest. Thus GWAS stands as just one of many tools that a genetics researcher must employ in the search for phenotype-associated SNPs.

These challenges aside, GWAS has been a powerful tool for colorectal cancer discovery, particularly in the search for predisposition SNPs. As of 2012, 20 CRC susceptibility loci have been identified, all using the methods of GWAS and imputation-based fine mapping (27; 59; 89; 90; 102; 195; 196; 197). Genome-wide study of 5-FU toxicity has been less successful to date, with only one GWAS attempt, hindered largely by sample size, but potentially also as a result of confounders due to pooling of patients on both 5-FU infusional monotherapy and FOLFOX (63). Finally, GWAS of cancer prognosis has only been tried for breast cancer, for which it did not identify any significant markers (discovery n=1,145) (11).

1.4.4 Survival Analysis

Survival analyses use time-to-event data and thus require different statistical approaches than do traditional association tests. Whereas continuous, quantitative or binary outcomes can be described in a single variable, time-to-event data require two variables: the length of time that a subject was followed (continuous) and whether this subject exited the study due to an event or due to censoring (binary). The simplest form of a survival analysis is to compare by logrank tests whether a difference exists between the proportion of cases remaining alive at each timepoint and the proportion of controls remaining alive at each endpoint, known as Kaplan-Meier curves (111). A more robust approach is the Cox proportional hazards method (42), which allows for the use of covariates by assuming that covariates are multiplicative to the inherent risk of death experienced by the cohort (the "baseline hazard"), and thus not bothering to measure it. Conversely, Kaplan-Meier analyses must be stratified by each covariate. Given the assumptions of Cox, it is important to test whether covariates do in fact have a consistent multiplicative effect, as can be achieved with Schoenfeld residuals (172).

Both the Kaplan-Meier and Cox methods can be applied to the study of genetic polymorphisms – the Kaplan-Meier method stratifies cases and controls based on genotype and can identify significance but not effect size, while the Cox method assumes that allele dosage has a multiplicative effect on survival and produces a beta estimate in either univariate or multivariate analysis. Cox proportional hazards has been the preferred model for most genetic analyses (11; 94; 177), given its ability to adjust for covariates and produce an effect size estimate. However, these multivariate results are sometimes reproduced with the univariate Kaplan-Meier curves so as to provide a visual demonstration of the effect over time, or in the case that proportional hazards does not hold.

1.5 Thesis Outline

In summary, for biomarkers of 5-FU toxicity, the combination of previous association testing and functional evidence provides only three markers of probable impact on toxicity. These three markers are all within *DPYD* (*2A, 2846T>A (D949V), and 1679T>G (I560S)), and even these have not been tested for an effect in patients taking capecitabine, often the first line therapy of choice. Further, it remains unclear whether evidence of an effect in one 5-FU regimen can be generalised to other 5-FU regimens, and identification of genetic biomarkers thus far has been restricted to a limited number of candidate genes, if not just candidate polymorphisms. Markers of CRC survival have been far less reported on, possibly the result of a failure to report null results, but also because the data needed to identify these markers are more complex, requiring long term follow up, with careful patient recruitment, for many patients, probably as part of a clinical trial.

To expand the field, this thesis first presents three chapters on 5-FU toxicity. First, a systematic review and meta-analysis of the decade of publications on genetic markers of 5-FU toxicity is provided, alongside a report of the effect of these markers in the QUASAR2 trial of capecitabine with and without bevacizumab, run from Oxford University. While this chapter provides a summary of the field to date, it makes clear that the known markers provide insufficient predictive power when used as a companion diagnostic kit. Next is a chapter that searches for further markers of 5-FU toxicity in and around 25 metabolic pathway genes, using microarray genotyping, imputation and next generation exonic sequencing of these candidate genes to capture both common and rare alleles. The over/under abundance of associations in these pathways is also assessed, and in the third results chapter, the SNP heritability of capecitabine-based toxicity is analysed prior to performing a two-study GWAS.

For CRC prognosis, a left truncation due to a delay in the collection of bloods from patients is identified in the data from the VICTOR trial of rofecoxib (Vioxx®,

Merck). Standard GWAS methods are first applied using Cox proportional hazards, and subsequently multiple methods are created that aim to improve statistical power, such as by incorporating data from other trials and from databases.

In using the germline to investigate two questions related to colorectal cancer treatment, this thesis applies multiple genetics approaches with the aims of improving personalisation of colorectal cancer treatment and expanding our research methods in the field.

Chapter 2

Materials and Methods

2.1 Samples Used

Samples from two Oxford-run clinical trials and one collaborative study group were used for primary analyses in this thesis, while further samples were used for validation.

QUASAR2

The QUASAR2 study is a multicentre international phase III randomised controlled trial of capecitabine (Xeloda®, Roche) with and without bevacizumab (Avastin®, Roche/Genentech) for adjuvant treatment of stage II/III colorectal cancer patients. As of July 2010 or August 2012 (the two dates for which patients were selected for array-based genotyping), 1,066 of the recruited patients had completed their chemotherapy and provided consent and blood for research. Various analyses in this thesis use different subsets of these patients, based on DNA availability at the time of genotyping and clinical data availability at the time of analysis. 987 of these patients were genotyped by Dr Claire Palles, Dr Enric Domingo and Dr Axel Walther using one of three Illumina genome-wide SNP genotyping panels: HumanHap 370 CNV (506 patients), HumanHap 610 (385 patients) or HumanOmni 2.5 (96 patients). Dr Palles

performed genotyping quality control (QC) measures (maximum 95% per-patient and per-SNP missingness, HWE p-value>1e-06, no sex mismatches with manifest, cryptic relatedness pi-hat>0.1875) and a principal components analysis of all patients, from which 946 passed, providing the quality control measures for the array-based genotyping in the meta-analysis (927 patients, after I excluded 19 for lack of clinical toxicity data available as of January 2012). Call rates for these patients were 99.83 (Omni 2.5), 99.94 (Hap610) and 99.85 (hap370). The mean SNP call rate post QC was >99.8 on each array. For the candidate-based, genome-wide and heritability analyses, I re-performed principal components for 940 included patients (available with clinical data received December 2012), all of whom remained in the analysis. These included 540 males, 400 females; 836 with colon cancer, 104 with rectal cancer; 337 with stage II disease, 603 with stage III disease; 439 receiving capecitabine monotherapy, 501 receiving bevacizumab in addition. Additionally, a largely overlapping 1,247 samples from QUASAR2 and CORGI (including intentional duplicates) were genotyped on Illumina HumanExome arrays by genomics facilities in Edinburgh and Oxford, for which I applied genotyping quality control measures and principal components analysis. These samples included 987 QUASAR2 patients with clinical toxicity data available as of December 2012. The exome array QC is detailed in Section 4.3. Further genotyping was performed by Dr Palles and me for various SNPs using KASPar (Section 2.2). For all QUASAR2 and VICTOR arrays, base calling was performed using Illumina Genome Studio (re-performed with ZCall for exome arrays), applying an 0.4 GenTrain and -0.25 no-call cutoff, and excluding SNPs with a het-excess score of <-0.3 or >0.2. Cluster plots were examined for a haphazard selection of SNPs, as well as for SNPs with associations reported in this thesis.

VICTOR

The VICTOR trial tested whether use of the COX-2 inhibitor rofecoxib (Vioxx®, Merck) post-resection reduced recurrence of non-metastatic colorectal cancer relative to placebo. Of the 2,434 patients entered into the study, only a proportion received treatment (mean duration of 7.4 months; only 18 patients completed the 2 year study treatment), as a result of the drug's worldwide withdrawal in September 2004 (142). Despite the terminated drug use, the trial serves as a valuable resource for genetic study of colorectal cancer, given the long term survival data available. In particular, 969 patients had bloods collected and were genotyped on HumanHap300 and HumanHap370 arrays by Dr Axel Walther and Dr Enric Domingo, of which 934 passed genotyping quality control performed by Dr Claire Palles (a current member of the Tomlinson Lab). I performed principal components analysis as part of the heritability analysis (Section 6.5). Using the final cleaned trial data (provided January 2012 by Sue Dutton of the Centre for Statistics in Medicine, Oxford), for the survival analysis, I removed an additional 16 patients with problematic data (survival dates missing or negative at-risk periods), leaving 918 patients (595 males, 323 females; 600 colon, 318 rectal; 464 stage I/II, 454 stage III; 572 receiving adjuvant chemotherapy, 346 receiving surgery only). Of these, 108 experienced an overall survival event and 189 experienced a disease free survival event (91 patients included in both measures).

CORGI

The CORGI (Colorectal Tumour Gene Identification) study aims to identify CRC predisposition alleles through multi-centre recruitment of CRC cases and controls UK-wide. 485 CRC cases from CORGI, for whom genotypes generated previously using the Illumina Hap550 array were available, are used in Section 6.7.2, as described there.

EPICOLON

85 capecitabine-treated patients from the EPICOLON study (1) were used as replication in Chapter 4. Genome-amplified DNAs were provided, and Dr Palles and I performed genotyping by KASPar.

SCP

The Spanish Capecitabine Pharmacogenetic (SCP) Study consists of 233 colorectal and breast cancer patients recruited prospectively through multiple oncology units in Spain, particularly Hospital Universitario San Carlos and Hospital Universitario Virgen de la Victoria. Colorectal cancer patients were treated with a standard regimen that matched that of QUASAR2 (standard: 1250 mg/m^2 twice daily d1-14 every 3 weeks), while breast cancer patients received either this regimen or a continuous regimen (continuous: 800 mg/m^2 orally every 12 hours daily). Toxicity data were collected according to the NCI's CTCAEv3, for the phenotypes HFS and diarrhoea (summarized in Chapter 4). Tumour type, age, gender and capecitabine regimen were recorded for each patient. gDNA was extracted from peripheral blood lymphoctyes, genotyped using Illumina Hap610 SNP arrays, and subjected to standard genomic quality control measures including per-patient and per-SNP call rates below 95% and HWE deviations of p<1e-06, with base calling performed in GenomeStudio and cluster plots checked manually for reported-on SNPs. Concordance among duplicate samples was >99%. Imputation and statistical analyses (with SNPTEST) were performed by Dr Pita Guillermo of the Spanish National Cancer Research Centre (CNIO), with oversight from Dr Anna Gonzalez-Neira, according to the same methods described in this thesis.

2.2 General Laboratory Methods

Consumables

Sources for most consumables are named within the specific method for which they are used. Consumables common to multiple methods are named here.

Water

Water for PCR was purified on site by a Purelab flex machine (ELGA). DEPC treated water (Ambion) was used for 454 experiments and for normalising DNA samples.

Agarose Gels

1% and 2% agarose gels were prepared from powder (Sigma) and 1x tris acetate ethylenediaminetetraacetic acid (TAE) buffer (VWR). To make one liter, 20 grams of powder were added to 1 liter of TAE and microwaved for 8 minutes, stirring regularly. After cooling, 20μl of ethidium bromide (Electran) was added. 5x Orange G (Sigma) loading dye (50ml 10xTBE, 80g sucrose, 100mg Orange G dye; stir 30 minutes; fill to 200ml with water) was used to weight DNA solutions and track movement into the wells and through the gel. Electrophoresis was performed at 120 volts for 20 to 30 minutes. Movement in the gel was tracked with 1kb Plus DNA Ladder (Invitrogen), made from 1ml Orange G, 4ml of 1xTBE and 200μl of stock in a 7ml tube.

DNA Extraction

From Blood

Only gDNA from human blood was used for experiments described in this thesis. gDNA extractions were performed using the Promega Maxwell 16 DNA Purification Kit: Blood DNA Purification Kit (Cat. AS1010), using a Maxwell 16 Instrument.

The Maxwell procedure uses paramagnetic particles to perform a mobile solid phase capture, wash and elution of gDNA from blood. Up to 16 samples can be run at once.

Quantification

Following extraction, concentration was quantified using either NanoDrop® (Thermo Scientific) or PicoGreen® (Invitrogen). NanoDrop® is a spectrophotometer to which 1μl of aqueous DNA is added. PicoGreen® is a dye that fluoresces 1000-fold upon binding of double stranded DNA (179). DNA concentrations are measured by extrapolation from a standard curve generated using stock DNA of known concentrations/dilutions. For this method, 1μl of aqueous DNA is diluted in 99μl of 1:20 TE buffer and mixed with 100μl of 1:200 reagent. PCR products can also be quantified using these methods. For the Roche/454 sequencing project described below, all samples, including those used to generate the standard curve, were measured in duplicate.

DNA Amplification

PCR

PCR was used to create amplicons from gDNA. Standard PCR reagents and their concentrations are provided in Table 2.1. Reagents and *taq* were either from Bioline or Roche. PCR reagents were generally combined in a master mix that was distributed into individual wells of a skirted 96-well plate (Thermo Scientific). PCR primers (Sigma) were generally 20-25bp in length and formed a 200-1000bp product, including a 25-50bp flanking region around to the target. PCR primers were designed with primer3 (`http://frodo.wi.mit.edu/primer3/`) using DNA sequence of the target and flanking region downloaded from UCSC Bioinformatics (`http://genome.ucsc.edu/cgi-bin/hgGateway`). Reverse complement DNA sequence was generated using Bioinformatics.org (`http://www.bioinformatics.org/sms/rev_comp.html`).

Primers arrived dehydrated and were dissolved in water to a concentration of 100μM, then diluted further to 10μM working dilutions for use in PCR. Amplification was performed on a G-Storm thermal cycler. Prior to amplification, DNA and master mix were spun down using a Megafuge 16R centrifuge (Thermo Scientific). PCR reactions were tested at $55 - 65\,^{\circ}$C annealing temperature to determine optimum conditions. Negative (no template) controls were included in each plate of PCR reactions.

Table 2.1: Volume of each reagent included in a standard single-sample 25μl PCR reaction.

Reagent	**μl**
10x buffer	2.5
$MgCl_2$ (25 mM)	2
dNTPs	2
Fw primer	0.5
Rv primer	0.5
Taq	0.25
DNA	1
Water	16.25
	25μl

Primer Optimisation

Different DNA regions amplify differently due to G/C content. By tweaking the above PCR protocol as per the troubleshooting steps described in Table 2.2, most regions can be amplified successfully. Generally, the first optimization variable tested was annealing temperature. Higher annealing temperature increases specificity because a perfect primer/template complement is required to overcome the primer/template melting point; however, shorter primers and primers that bind to a region that includes a variable base may only bind at lower temperatures. Amplicons that failed to amplify despite troubleshooting were re-designed or amplified by joining adjacent primer pairs.

Table 2.2: Altering the temperature cycles and/or reagent concentration can improve amplification levels and specificity.

Amplification Problem	**Solution**
Strong primer dimer band in addition to desired PCR product	Decrease concentration of primer so less is left over.
Non-specific product (multiple bands and/or a highly streaky lane)	Select best annealing temperature, then re-test at this temperature using less $MgCl_2$(1.5μl) and less primer (0.25μl forward and reverse total).
Amplification failed at all annealing temperatures	Re-test with additional cycles and lower $MgCl_2$. If this fails, re-design primers.
Band is weak	Raise $MgCl_2$ concentration (2.5mM) and increase the number of cycles (eg, 40).
Band is weak despite higher $MgCl_2$ and additional cycles	Add DMSO and increase the extension time.

Purification

Post-PCR reactions contain both the desired amplicons and many impurities, including left over PCR reagents (dNTPs, buffer, etc.), primers, early terminated amplicons, and, in some cases, undesired full length amplicons. Other than unwanted full length amplicons, which have the same chemistry as full length desired products, performing a PCR library purification eliminates PCR contaminants. PCR purification was performed using the Agencourt AMPure [XP] system (Beckman Coulter Genomics), which recovers amplicons of 100bp or higher by binding DNA amplicons to magnetic beads, separating these beads with an external magnet, removing impurities with ethanol, and then eluting the DNA from the beads and transferring DNA to a new PCR plate (`http://www.beckmangenomics.com/products/dna_purification_and_cleanup/agencourt_ampure_xp.html`).

Genotyping

KASPar

KASPar (KBiosciences) is a competitive allele-specific genotyping method, run in 96-well plates on a real-time PCR instrument (Applied Biosystems HT7900). This method discriminates between alleles by having separate fluorescent primers for each allele compete for binding of the DNA and subsequent proliferation (`http://www.kbioscience.co.uk/reagents/KB_AN_KASPar_chemistry.pdf`). Primers for the assays were designed using Primer Picker (KBiosciences). Two fluorescently labelled primers were designed that differed only at the position of the SNP – the 3'-most nucleotide of the primer. These primers amplify the DNA and increase the intensity of a VIC or FAM label. Bi-allelic SNPs can be genotyped with this method, the output of which is an X/Y cluster plot in which wild-type samples have high Y-intensities and low X-intensities, homozygous recessive samples have high X-intensities and low Y-intensities, and heterozygote samples cluster in the middle. Table 2.3 provides reagents and respective concentrations used in a standard KASPar mastermix, though the components of the reaction mix are not divulged by the company.

Table 2.3: Volume of each reagent included in a standard single-sample 10μl KASPar reaction. 100μl of assay mix contains a forward primer for each of the two SNP alleles (12μl each), a reverse primer on the non-polymorphic site (30μl) and water (46μl).

Reagent	**μl**
2x Rxn mix	5
Assay mix	0.14
DNA	1
Water	3.86
	10μl

Sanger Sequencing

The original wide-spread sequencing method, named "Sanger sequencing" for its inventor, utilises fluorescing chain-terminating nucleotides that stop DNA elongation at the locus at which they are incorporated (170). Used at the correct concentrations alongside normal nucleotides, these modified nucleotides will halt DNA elongation at each locus, allowing researchers to obtain a DNA sequence by using gel electrophoresis that separates DNA strands based on size. Today, this method is used primarily for small scale sequencing (ie, few loci to sequence), to sequence long continuous stretches of DNA (500-1000bp) and as technical validation of next-generation sequencing. In this thesis, Sanger sequencing is used primarily for exon sequencing of one gene. Table 2.4 provides reagents and respective concentrations used in a standard Sanger sequencing mastermix, which is used after amplifying by PCR and cleaning the PCR product using ExoSAP (Affymetrix). The sequencing reaction runs for about 4 hours in a thermocycler.

Table 2.4: Volume of each reagent included in a standard single-sample 20μl Sanger sequencing reaction. BDT=big dye terminator (Applied Biosystems). Often duplicate reactions are performed for a given sequence, one in the forward direction, one in the reverse direction. Primer concentration is 10mM.

Reagent	**μl**
Cleaned PCR product	1
Fw or Rv primer	0.5
BDT	8
Water	10.5
	20μl

Microarrays

DNA microarrays allow for the assessment of hundreds of thousands of loci simultaneously in a single array experiment, reducing experimental error and cost. In the early 2000s, this technology was first applied to gene expression levels, and by the mid-2000s

it was extended to SNP genotyping (78). The method relies on haplotype blocks to allow hundreds of thousands of tagging SNPs to capture much of the genome-wide diversity (106). Briefly, whole genome DNA is amplified, and then for each genotype target locus, target and surrounding gDNA is captured by site-specific beads to which 50-mer complementary probes are attached that only become fluorescently labelled when binding a particular allele (78). This, in turn, is read photographically by a bead array scanner for thousands of beads at once.

Most of the analysis described in this thesis is of tagging SNP (Illumina HumanHap300 duo-, HumanHap370 CNV duo-, HumanHap610 quad-, and HumanOmni2.5 quad-bead chips) and exonic SNP (Illumina HumanExome 12-bead chip pool v1 and v1.1) microarray data. While I prepared some of the DNA samples for run on exome arrays, the bulk of the microarray data generation was performed by others prior to when I started my research, namely Dr Axel Walther, Dr Claire Palles and Dr Enric Domingo. Given the economies of scale related to technical quality control measures, microarrays are increasingly run in specialised genomics centres. As such, exome array genotypes were generated at the Genomics Centres in Oxford and Edinburgh. Genotyping quality control measures and analysis are described throughout this thesis as they apply to particular projects (in particular, see: Section 4.3, Section 5.4, Section 6.6, Section 6.7.2, and Section 6.8).

Next-Generation Sequencing

Since 2005, various DNA sequencing methods have existed that do not rely on Sanger-like capillary methods (137). Collectively, these "next-generation" methods allow researchers to sequence entire genes, exomes and/or genomes of humans and other species at increasingly affordable cost. Of these next-generation methods, this thesis uses the Roche/454 Titanium GS FLX next generation sequencing platform for amplicon sequencing, as described in Section 2.3.

2.3 Roche/454 Next-Generation Sequencing

Overview

Chapter 4 employs Roche/454 Titanium GS FLX next generation sequencing technology in search of functional, and likely rare, variants tagged by various SNPs identified within or near to three 5-FU/capecitabine pathway genes: *DPYD*, *TYMS* and *UPP2*. Identification of functional alleles took the form of identifying a pooled frequency difference between 200 patients from the QUASAR2 trial whose DNA was sequenced in one of two groups, a high toxicity ("HT") pool of the 100 patients experiencing the most toxicity and a low toxicity ("LT") pool of the 100 patients experiencing the least toxicity. Critical to this method was the accurate calling of rare variants and the precise pooling of all patients' DNA in equal concentrations. We aimed to reach 3000-read coverage per pool at each target locus, in order to surpass 25-read average coverage per patient.

The specific sequencing methodology (eg, primer design and amplicon pooling strategy) is described here, while the Technical Results section of Chapter 4 (Section 4.4) show the coverage achieved and variant filtering performed, and the effect of these variants on 5-FU toxicity are presented in Section 4.5.9. For this project, Dr Palles and I jointly prepared amplicons for sequencing and these were sequenced by the BRC Group at the Wellcome Trust Centre for Human Genetics, Oxford. Largely, our method was adapted from a Roche protocol: Application Note 5 mplicon Sequencing, February 2007 (`http://www.454.com/downloads/protocols/5_AmpliconSequencing.pdf`).

Technology

Sequencing by Roche/454 (so called because the company Roche owns the company Life Sciences, of which 454 is a subsidiary) technology requires four steps: library

preparation, DNA amplification, sequencing, and signal processing, as described by Balzer et al (15) and 454 Life Sciences.

In library preparation, a nebulizer is used to break double stranded gDNA into fragments of length 400-600bp, to which enzymes attach adapters that prime the sequence for both amplification and sequencing. These double stranded fragments are then separated into single stranded DNA fragments that are amplified via emulsion-based PCR. Critically, the bead-based sequencing employed later in the method requires that many copies of exactly one DNA fragment are bound to each bead. Emulsion PCR accomplishes this by performing PCR within water droplets formed around each bead, each droplet containing one DNA fragment prior to amplification. Pyrosequencing is then performed via a sequencing-by-synthesis method. Specifically, A, T, G, and C nucleotides are separately washed over a PicoTiterPlate™ that contains about one million wells, each containing at most one bead to which single-stranded DNA is bound. As each nucleotide is incorporated by complementary base pair matching, a luminescent chemical is formed. At the end of each wash round, a camera identifies both the colour (one for each nucleotide species) and intensity (one unit per nucleotide incorporated) produced by chemi-luminescence.

Advantages. Advantages to this technology include a massively parallel method that was novel in 2005, a read length of up to 500bp that still remains industry-leading among methods that have achieved low error rates (ie, not nanopore sequencing) and an ability to accurately call insertions and deletions. While average error rate for Roche/454 sequencing is about 1%, this rate falls to about 0.6% when excluding homopolymers.

Limitations. The main limitation of Roche/454 technology is that it can not accurately sequence non-complex stretches of DNA (ie, homopolymer tracts). Based on the method's technology, sequential bases incorporate and fluoresce in a single step, and while the sensors can accurately discriminate between the signal intensity

of one and two of the same base, it becomes less discriminate when the intensity corresponds to six or more bases (137).

Patient Pooling

DNA for the 200 patients selected for sequencing was extracted using Maxwell 16 kits (Section 2.2). DNA concentration was measured by PicoGreen (Section 2.2). The 100 low toxicity patients ("LT") were combined into 5 pools of DNA from 20 individuals each, as were the 100 high toxicity ("HT") patients. As shown in Figure 2.1, these 10 pools of DNA were the templates for PCR for each of the 49 amplicons across *DPYD*, *TYMS* and *UPP2* (amplicons described further in the next section). Pooling of DNA for library preparation was possible because the final step of the protocol required pooling all patients into two pools, one for the LT patients and one for the HT patients. Pooling the 200 patients into 10 pools greatly reduced reagent costs and time; amplicon preparation as one pool of HT DNAs and one pool of LT DNAs was theoretically possible, but would have increased the risk that amplicons from some patients were lost by chance.

	1	2	3	4	5	6	7	8	9	10	11	12
A	LT Pool 1	LT Pool 2	LT Pool 3	LT Pool 4	LT Pool 5	HT Pool 1	HT Pool 2	HT Pool 3	HT Pool 4	HT Pool 5	(-)	(+)
	Amplicon 1 fusion primers ----------|											
B	LT Pool 1	LT Pool 2	LT Pool 3	LT Pool 4	LT Pool 5	HT Pool 1	HT Pool 2	HT Pool 3	HT Pool 4	HT Pool 5	(-)	(+)
	Amplicon 2 fusion primers ----------|											
	...continued for multiple plates...											
X	LT Pool 1	LT Pool 2	LT Pool 3	LT Pool 4	LT Pool 5	HT Pool 1	HT Pool 2	HT Pool 3	HT Pool 4	HT Pool 5	(-)	(+)
	Amplicon 49 fusion primers ----------|											

Figure 2.1: 10 pools of gDNA from 20 LT or 20 HT patients each were formed in equal concentration prior to amplification of the 49 target amplicons, simplifying the protocol for certain steps such as quantification and concentration equalisation. In some steps, further pooling was performed; in the final step, two pools were formed, one of DNA from all HT patients and one of DNA from all LT patients.

Using these pools, first amplicons were formed by PCR, then amplicons were purified, quantified and equalised in concentration, then a final HT and a final LT pool of DNA from 100 patients each was formed for sequencing, as shown in the workflow diagram in Figure 2.2.

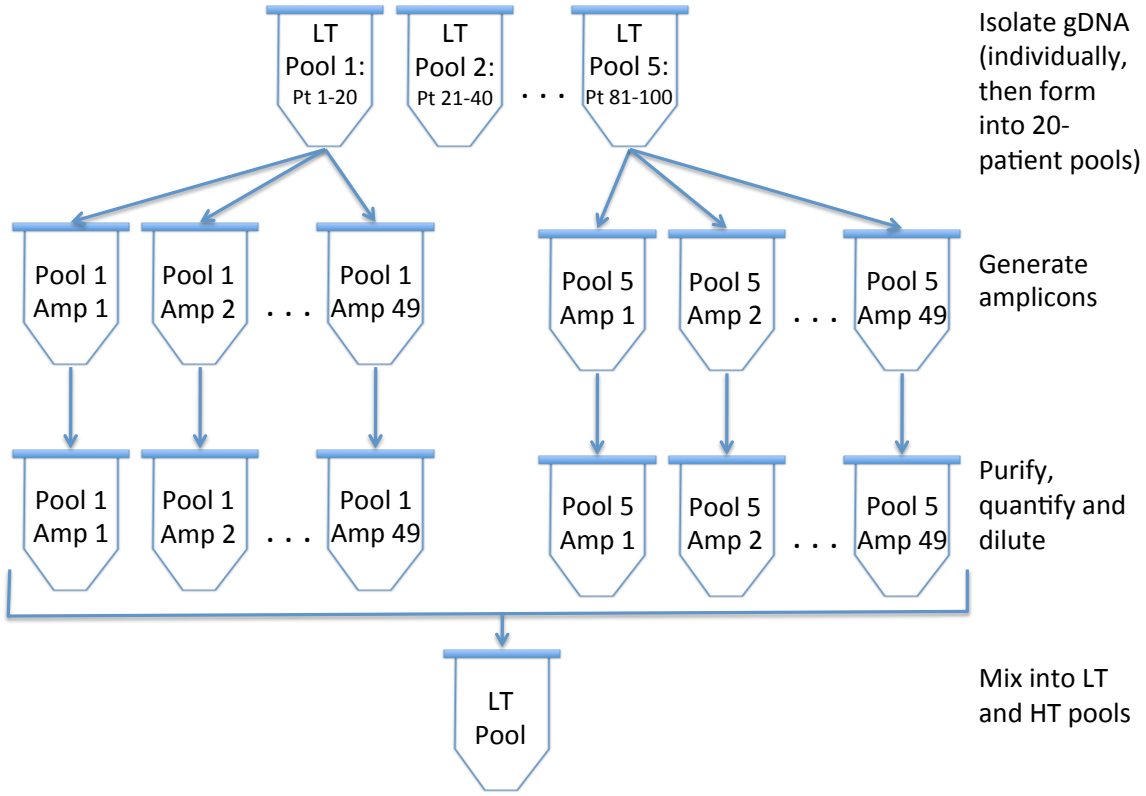


Figure 2.2: 4,900 amplicons from low toxicity patients and 4,900 amplicons from high toxicity patients were generated in pools of 20 patients, purified, quantified, and diluted, prior to forming one final low toxicity pool and one final high toxicity pool.

Importantly, LT and HT gDNA was kept separate throughout every step of the experiment. Contaminating LT DNA with HT DNA (and vice versa) would destroy the ability to identify variants associated with a low or high toxicity phenotype because each pool would contain variants associated with both low and high toxicity. Similarly, it was vital for DNA concentrations to be uniform across all patients. Because patient samples were pooled, variant frequency was determined as an average frequency for each variant among all patients in a given pool. Including extra DNA for a single patient would alter the frequency of the variant for that pool toward that

patient genotype, potentially making the LT and HT pools falsely more genetically similar or falsely more genetically different.

Primer Design and Optimisation

Following PCR optimisation using commercially available gDNA, 49 amplicons were generated using Roche FastStart HiFi reagents. These amplicons corresponded to the coding and regulatory (splice site, CpG and UTR) regions of *DPYD*, *TYMS* and *UPP2*. Negative and positive (commercial DNA) controls were included in each PCR.

Exon coordinates for *DPYD*, *TYMS* and *UPP2* were downloaded from the UCSC Bioinformatics website, Table Browser section. In total, coordinates for 39 exons were downloaded. Where CpG islands (*DPYD* exon 1 and *TMYS* exon 1) and UTRs existed, exon ranges were expanded manually to include the entire CpG island range, as identified by the UCSC Bioinformatics website, Genome Browser section. The exon, UTR, splice, and/or CpG island ranges formed the "target regions".

Sequence-specific primers were designed such that a 25bp flanking region was present for each target region, as shown in Figure 2.3. Table 2.5 and 2.6 provide the sequences of these primers, first for *DPYD* and then for *TYMS* and *UPP2*, as well as the conditions used in amplification. At most, one exon was captured per target region. Target regions of greater than 500bp (eg, *DPYD* exon 1) were broken into multiple amplicons of under 500bp (as is optimal for Roche/454 sequencing). In some cases, difficulties with amplification required amplicons to remain over 500bp. To minimise sequencing of non-target regions, primers were designed within a 100bp region flanking the target region, though closer to 25bp was preferred.

Initial sequence-specific primers shown in these tables were expanded into Roche/454 "fusion primers" when ordering. These primers contain a combination of sequence-specific primers and a fixed sequence tag – a 19-mer that (a) acts as a priming site during clonal amplification, (b) allows binding of DNA to capture beads and

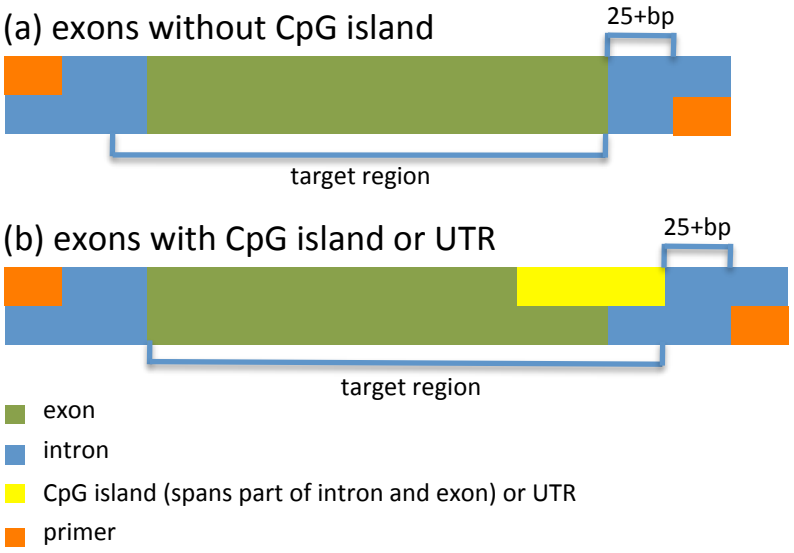


Figure 2.3: Roche/454 amplicon sequencing primer placement allowed for the amplification of either the exon and adjacent splice sites (a) or the exon, splice sites and additional regulatory regions (b).

(c) serves as a priming site for the sequencing reaction (Application Note 5 – Amplicon Sequencing, February 2007, `http://www.454.com/downloads/protocols/5_AmpliconSequencing.pdf`). Specifically, the primer A adaptor sequence (5 GCCTCCCTCGCGCCATCAG 3) was attached to the 5-prime end of each forward sequence specific primer and the primer B adaptor sequence (5 GCCTTGCCAGCCCGCTCAG 3) was attached to the 5-prime end of each reverse sequence specific primer. For example,

GCCTCCCTCGCGCCATCAGTGGAGTGAGGTAATTTTCAGCA

and

GCCTTGCCAGCCCGCTCAGTGCCAACATATTTCCATATATTTACAC

for *DPYD* exon 2 forward and reverse, respectively.

PCR conditions included constant annealing temperatures across all cycles. Unless noted elsewhere, extension time and annealing time were 30 seconds, extension temperature was 72 °C, 0.25μM primers and 2μl of 25ng/μl DNA were added per sample, and total volume was 25μl. PCR amplification was checked by 1% gel elec-

trophoresis for size and specific product. Once all amplicons had been optimized using Bioline *taq*, amplification was re-tested using Roche hot start *taq*, which was used to amplify the pooled DNA for sequencing. After amplification, DNA was purified and concentration of each amplicon was measured by PicoGreen for each patient pool.

DNA Equalisation

Once amplified and purified, it was important to quantify and equalize the concentration of DNA present for each of the 49 amplicons from each of the 10 patient pools. It was vital to combine into a final mix for sequencing an equal quantity of DNA from each patient because the number of 454-sequencing read counts of a given patient's DNA follows a normal distribution weighted by its percent of the total DNA being sequenced.

Each patient/amplicon combination was checked in duplicate by PicoGreen. The average concentration was used, and where concentrations differed substantially in the two measurements, a third measurement was taken. Amplicon concentration was then equalised by diluting samples to a consistent number (10^9) of DNA molecules per microliter. The number of molecules per microliter was calculated from concentration using Equation 2.1.

$$\text{Molecules}/\mu l = \frac{\text{sample conc [ng}/\mu l] \times 6.022 \times 10^{23}}{656.6 \times 10^9 \times \text{ amplicon length [bp]}} \quad (2.1)$$

The number of molecules per microliter present in each sample was altered to the consistent number of molecules per microliter (10^9) by adding 1 x TE buffer. The amount (in μl) of 1 x TE buffer to add to each plate, based on the number of molecules per microliter present, was calculated by Equation 2.2.

$$1 \times \text{TE} = \left(\frac{\text{molecules}/\mu l}{10^9} - 1 \right) \mu\text{l} \quad (2.2)$$

Target Sequence	Target Region (UCSC build 18)	Target Length	PCR Product(s) Length	Primers Without Adaptor (Forward; Reverse)	Annealing Temp. (°C)	Cycles	$MgCl_2$ (µl)	Notes
DPYD Exon 1 and CpG	chr1:98158778-98159312	534	385	CTCTGGAACTCCAGTGCTTCTT; AGACTCGAGACTGTAGGCACTG	59	35	1.5	1 min. extension
			547	GAAACTTTCCCGCGTCTCTC; CGGACTGCTTTTACCTTTGC	63	40	1.5	1 min. extension; 2.5µl DMSO
DPYD Exon 2	chr1:98121408-98121518	110	206	TGGAGTGAGGTAATTTTCAGCA; TGCCAACATATTTCCATATATTTACAC	55	35	2	
DPYD Exon 3	chr1:98066258-98066340	82	151	TGTTGCTATGGTGACCCAGA; TCAGGATCTTAGAGAATTAATGCAA	59	35	2	
DPYD Exon 4	chr1:97978536-97978623	87	221	CAAATACTGTTATTTTCATTTGCAG; ACTAGACACGGACTCTGAATGAG	57	40	1.5	1 min. extension; 1 min. annealing
DPYD Exon 5	chr1:97959654-97959815	161	279	GTGCATGGTGATGGTAGTGG; TTTGTCGTAATTTGGCTGTTT	63	35	2.5	
DPYD Exon 8	chr1:97937495-97937691	196	269	TTTGCTCCATCATTTCTGACA; TGACAATTGATTTCCCCGTA	63	35	2.5	
DPYD Exon 9	chr1:97929861-97929942	81	200	GCTTACTCCTTTCTTTTTGAGCAG; TTAAGATTGGTCAAAGATTGGTCA	63	35	2.5	0.25µl primers
DPYD Exon 10	chr1:97917239-97917326	87	255	GCAGTCATTCTTCTGGATATTGC; GGCTGACTTTTCATTCTTTTTGA	57	40	2	
DPYD Exon 11	chr1:97833203-97833310	107	235	TGCTGCTGAGCTTGATTTTG; CCCCTCCTCCTGCTAATGA	63	35	2.5	
DPYD Exon 12	chr1:97831362-97831531	169	221	CCATATTTTCTGATGGTTCCATTT; TGGGGTTTTCCTTTCATCAT	63	40	2	
DPYD Exon 13	chr1:97811904-97812114	210	431	ACATTTCAGCTCCCAGCACT; CTCCAATTCAAGACTCCCTGA	59	35	2	
DPYD Exon 14	chr1:97787704-97787888	184	263	AAGTGGAAATGATGGCAAATG; GGAATTGTCTGTTGACTAGTGATTG	63	35	2	
DPYD Exon 15	chr1:97753870-97754085	215	443	AATGTGTAATGATAGGTCTTGTCAAA; TTCGGATGCTGTGTTGAAGT	63	35	2	
DPYD Exon 16	chr1:97688203-97688367	164	261	TCACCAACTTATGCCAATTCTC; GCTTTTCTTTGTCAAAAGGAGAC	63	35	2	
DPYD Exon 17	chr1:97620537-97620605	68	332	TTTCTCATGGCAGCTCTTTATTT; CAAATGTCATCCAGTGTTTTGC	63	40	2	
DPYD Exon 18	chr1:97611705-97611788	83	208	GGGGCCTGTTATGTCACACT; TCAACGGTGAAAGCCTATTG	57	35	2	
DPYD Exon 19	chr1:97544321-97544441	120	238	TTACAATAGCGGGCAACTGA; AGCTTTGCTGTTGTTCCAGT	59	35	1.5	1 min. extension
DPYD Exon 20	chr1:97543403-97543522	119	203	TCAGCAACCTCCAAGAAAGC; AAGAGCTGCATGAAAATGTTG	57	35	2	
DPYD Exon 21	chr1:97472996-97473138	142	240	CATTGCATTTGTGAGATGGAG; GTGGTCAGTGTGCTAACCATT	63	40	1.5	
DPYD Exon 22	chr1:97431213-97431392	179	267	CCAAGGCTGAGTTCTCAAGAATA; CTCTTGGTTTGACACTTGGTTG	57	35	2	
DPYD Exon 24	chr1:97336633-97336776	143	232	GGCATACTTATATAGTTGGCATAATTT; CCCATTTTTCTCTTCTCTGAGC	57	35	2	
DPYD Exon 25	chr1:97320474-97320614	140	251	TCTGCCGTAAAAACAAGAAGAA; GAAAATGTTTTAAAATGAGCTTGC	57	40	1.5	
DPYD Exon 26	chr1:97315887-97317290	1403	415	ATGAGGCAGCAAAAGACCAT; GCTCAAAGTATGCAATGCTTCA	65	35	2.5	
			512	TCAAACTGATTTGGCACACTT; TGTATCCAAGGAAACATTTTATCA	59	35	2	
			348	TGACTTTCTTCATTTGTACTTTATGGA; TCATGTCAATGACCATTCAATTAGT	59	35	2	
			477	AACTGGCAGTGAACATCCAA; GGTCAGTGACATCAATACCCTCT	59	35	2.5	

Table 2.5: 27 primer pairs were used to amplify the 23 target regions in *DPYD*. After multiple optimisation rounds, amplicons were generated using the conditions described in this table. Note that the PCR product length excludes the length of the 38bp adaptors.

Target Sequence	Target Region (UCSC build 18)	Target Length	PCR Product(s) Length	Primers Without Adaptor (Forward; Reverse)	Annealing Temp. (°C)	Cycles	$MgCl_2$ (µl)	Notes
TYMS Exon 1 and CpG	chr18:647402-648745	1343	632	AGGAAACAGATCTCAAACAGCA; GCGTCACCTCTCAGGCTGTA	55	35	2.5	1 min. extension; 2.5µl DMSO
			671	GCACCCTGTCGGTATTCG; CGGAAAGGAAAGAGGAAAGG	63	35	2	1 min. extension
			300	CACGAACCAGCTTTCCTCTT; ATCTGAGCGGTCAAGGTCAG	63	40	1.5	
TYMS Exon 2	chr18:649641-649714	73	250	GTCCCAGGGCAGTTTTCTTC; ACACCTGGCTGCTTTGAGTT	57	35	2	
TYMS Exon 3	chr18:652146-652320	174	311	TTTACCTGGATGCCTGCTCT; CATTGACTCATCCCCCAGAC	55	35	3	
TYMS Exon 4	chr18:659072-659173	101	247	TGACCTCTAAGGCCATCTCA; TCTGCACAAAACCAGATTGC	59	35	2	
TYMS Exon 5	chr18:660692-660867	175	260	CAAACCACCATCCCTCCTTA; AGTGTGGTAGGCTGGCAAGT	59	40	1.5	1 min. annealing and extension; 2.5µl DMSO
TYMS Exon 6	chr18:661380-661451	71	185	TTTCTCCCCCGGGTTTATAG; TCTTGGACACAGCATTGAGC	57	40	1.5	
TYMS Exon 7	chr18:662860-663499	639	446	TTGTTTCACGGACATGAGGA; CCTCAGCATTGTCAGATACCC	57	35	2	
			394	TTTTTAAGGATGTTGCCACTG; GGCAGTACAGAATGAACAAAGC	57	35	2	
UPP2 Exon 1	chr2:158559937-158560017	80	205	CACAATTCAGAGCCCAGTCA; GACTTGGAATATGACCAGACACA	65	35	2.5	
UPP2 Exon 2	chr2:158580559-158580644	85	391	TTATATACGGATGTGGTTTTGCTC; GCGTTAGTCCTTTTTCATACCACT	65	40	2.5	
UPP2 Exon 3	chr2:158666628-158666883	255	357	TGAGAGCTGGGCTGTGGTAT; ATCCAAGGAGCAACCATCTG	59	35	3	0.25µl primers
UPP2 Exon 4	chr2:158670857-158670974	117	240	CCCTAAAGGAAGGTTTATGTTCA; TCATTTTTCCTGAAATTAGAAATGT	59	35	2	
UPP2 Exon 5	chr2:158679859-158680017	158	232	CTATGGCAGGCCCTTGTTTA; GCCTGCAATGACTAGCTGAA	57	35	2	0.25µl primers
UPP2 Exon 6	chr2:158682582-158682696	114	218	TCACTTTGCAAGATGTATGATGA; GGCAGCTACCATGTATGTAAGG	57	35	2	
UPP2 Exon 7	chr2:158686167-158686376	209	410	TGTGTCAAAAGTAAACTCTAAACGA; GCTGGCATTTTTCTGTTTCC	55	35	2	
UPP2 Exon 8	chr2:158688507-158688653	146	219	GGGTGGGACATCAAGTCACTA; TTGACCCTAAGCATTCACGA	59	35	2.5	
UPP2 Exon 9	chr2:158699506-158700912	1406	424	GCTAGGGATGCTGGTGTGTT; TGAAATTTGAATTTAATTTAGTGTGAG	57	35	2	
			447	TGAAATCCTTCTCTCTTAAAAAGG; CAAACAGGCATATGTAAATCTCG	57	35	2.5	
			460	CTGCACTATTTCAGTGCACCA; CATCCAGTCATTCCGACTCC	57	35	2.5	
			374	ACTGATCGCATCTTGGGTTC; GCCAAAAAGAAGTTCACAAACC	59	35	2	

Table 2.6: 10 primer pairs were used to amplify the seven target regions in *TYMS* and 12 primer pairs were used to amplify the nine target regions in *UPP2*. After multiple optimisation rounds, amplicons were generated using the conditions described in this table. Note that the PCR product length excludes the length of the 38bp adaptors.

Quality Control of Amplicon Pools

Before forming the final two pools (high toxicity and low toxicity amplicon pools) for sequencing, it was important to confirm that only the desired products were present. Agilent High Sensitivity DNA Kits contain chips and reagents that allow for sizing and analysis of DNA fragments via electrophoretic separation of nucleic acids in microchannels. These kits were used to confirm the following characteristics of our amplicon pools (at this point, pooled in 4 LT pools and 4 HT pools, in order to fit all of our amplicons on the 11 wells of a single chip, while also spreading out amplicons of similar size in order to better visualise and measure each amplicon; technical results in Section 4.4):

1. Confirm the presence of each desired amplicon.
2. Confirm the absence of non-desired amplicons.
3. Confirm the absence of primer dimers.

Agilent high sensitivity DNA Kits were used in an Agilent 2100 Bioanalyzer, according to the standard Agilent High Sensitivity Assay Protocol (`http://www.genomics.agilent.com/files/ManualG2938-90320_HighSensitivityDNAQuick.pdf`) for each of the eight pools. When pools either appeared contaminated (non-desired amplicons or primer dimers) or lacked desired peaks, pools were re-made according to the previously described amplification, purification, quantification, and equalization protocols.

Sequencing

As specified in the amplicon library preparation protocol, for sequencing of each pool, we needed to provide 200μl of DNA/water solution containing 10^9 molecules per microliter. For the final pooling step, all of the amplified PCR product from all

of the pools used in the Agilent Bioanalyzer step were mixed into two falcon tubes (one for high toxicity, one for low toxicity), centrifuged for 30 seconds at low rpm, mixed by vortex for 1 minute, centrifuged again, then vortexed again, ensuring that all amplicons would be evenly distributed in the solution. Once mixed, for each of the two pools, 2.23μl of amplicon solution was added to 1.97μl of DEPC treated water and given to the BRC Group for Roche/454 sequencing.

Sequencing was performed by Dr Alistar Pagnamenta in a collaborating group, the BRC Group of the Wellcome Trust Centre for Human Genetics. Following analysis of sequencing results, amplicons with fewer than 3,000 reads at all target loci were formed into a new sequencing pool and re-sequenced and the data files were merged.

Mapping

Mapping and variant calling were predominately performed using two pieces of Roche-created software: GS Reference Mapper for mapping and GS Amplicon Variant Analyzer (AVA) for variant calling. In addition, UCSC Genome Browser and Microsoft Excel were used to aid in the analysis of sequencing coverage and read count. Mapping was performed by Dr Kulvinder Kaur of the BRC Group.

Between mapping and variant calling, adequate coverage (at least 30 reads per locus per patient) was checked using bed files obtained from GS Mapper. Where duplicate read counts existed per locus, the larger read count was used, given that duplicates resulted from early-terminating reads. Bed files were checked numerically in Excel for each of the 30 target regions and by creating UCSC Genome Browser custom tracks, build hg18/36, dense pack.

Variant Calling: Filters Based on Sequencing Statistics

Once sufficient read coverage was confirmed for all amplicons, the GS Amplicon Variant Analyzer (AVA) program was used to provide a list of potentially real (ie, not a

sequencing artifact) polymorphisms. In theory, none of the variants in the sequencing data should have appeared with below 0.5% MAF, which would occur when one of the 100 patients in a toxicity pool held the minor allele on one of their two chromosomes. In order to avoid creating false negatives (mistakenly removing true variants), the minimum MAF filter was set to 0%, which allowed for the appearance of rare variants that may have been underrepresented in the sequencing data due to low gDNA or amplicon concentration. Likewise, the maximum MAF filter was set to 100%.

The following manually-created filters were then applied to individual (ie, not "consensus") reads of the automatically-called variants, in order to remove further sequencing artifacts based on sequencing statistics:

1. **Total read count**: remove any variant whose locus has fewer than 3,000 reads of sequencing coverage in total between the forward and reverse direction, for each of the LT and HT pools.

2. **Bi-directional read count**: remove any remaining variant that has fewer than 50 reads in either the forward or reverse direction for either the LT or HT pool.

3. **Forward/reverse direction allele frequency discrepancy** (within each individual pool – HT and LT):

    - Rare variants: keep any variant whose forward and reverse direction allele frequency were below 0.1% (ie, rare variants assumed to be real at this stage).
    - Infrequent variants: if both directions show less than 5% frequency (but at least one is above 0.1%), remove a variant if the nominal difference in MAF between forward and reverse directions is greater than 2 (eg, remove if forward direction MAF = 1% and reverse MAF = 3.3%).
    - Common variants: if one direction shows equal to or greater than 5% frequency, remove a variant if the percent difference in MAF, measured

from the greater of the two MAFs, is greater than 75 (eg, keep if forward MAF = 15% and reverse MAF = 5%, but remove if forward MAF = 21% and reverse MAF = 5%).

We tried to create exclusion rules more strictly grounded in statistics, but found them overly stringent, and we feared the automatic exclusion of variants that would appear real upon manual inspection. This was partially expected, given that many filtering schemes merely require a single read in each direction (per patient). Instead, we guided our filtering method by using known variants, incrementally increasing the stringency of our filters until we removed dbSNPs that we knew would be polymorphic in our data from previously generated SNP array data. The use of dbSNPs allowed us to calibrate our method to high sensitivity (limiting our false negatives – those true variants that were called artifacts), but not to high specificity (limiting our false positives – those artifacts that were called true variants); for example, applying no filtering would have resulted in 100% sensitivity but 0% specificity. Without an equivalent metric for specificity, we needed to apply further stringent filtering in order to ensure that sequencing artifacts were sufficiently removed from our data.

Variant Calling: Individual Inspection

To improve specificity, the raw read data were visualised using AVA. This was performed for all variants that remained after applying the manual filters. Five criteria were selected as clear indicators of false variant identification, based on the known shortfalls of the Roche/454 method:

1. **Homopolymer**: variant falls within a stretch of 5 or more nucleotides of the same letter:
   AAAA[A/C]ACCAA

2. **Run of consecutive SNPs**: variant appears false because it falls within a

stretch of sequence with many called polymorphisms:

AA[A/C]T[A/T]CC[AA/TC]GC

3. **Early read termination**: variant is only observed near to the end of a sequencing read that has terminated before the end of the amplicon, and thus the sequencing software has trouble with sequence alignment:

   read 1 GGGTCTAAATTGCATTATAGGCGTAGGA

   read 2 GGGTCTAAATTGCATTATAGGCGTAGGA

   read 3 GGGTCTAAATTGCAT[T/A]

   read 4 GGGTCTAAATTGCATTATAGGCGTAGGA

   read 5 GGGTCTAAATTGCAT[T/A]

4. **Proximity to amplicon end**: variant is near the end of an amplicon, and as such is near the end of a sequencing read, for which the sequencing software has trouble with sequence alignment:

   read 1 GGGTCTAAATTGCATTATAGGCGTAGGA

   read 2 GGGTCTAAATTGCATTATAG[G/C]CGTAGGA

   read 3 GGGTCTAAATTGCATTATAGGCGTA[G/A]GA

   read 4 GGGTCTAAATTGCATTATAGGCGTAGGA

   read 5 GGGTCTAAATTGCATTATAGGCGTAG[G/A]A

5. **Unusually high sequencing noise**: many variants are present in the area in an inconsistent fashion and the origin of the sequencing disturbance is unclear:

   read 1 GGGTCTAAAT[T/G]GCATTATA[G/C]GCGT

   read 2 GGGTCTA[A/A]ATTG[C/A]ATTATA[G/C]GCGT

   read 3 GGGTCTAAATTGCATTATAGGCGTA[G/A]GA

   read 4 GGG[T/A]CTAAA[A/T]TGCA[A/T]TATAGGCGT

   read 5 GGGTCTAA[A/C]TTGCATTA[T/A]AGGCGT

As an example, Figure 2.4 shows a locus (highlighted in blue) that was falsely identified as an A/T variant because it falls within a homopolymer. Adjacent to this falsely identified A/T SNP is a falsely identified A insertion. The homopolymer in this example is quite long – 7 consecutive Ts. At this high count of Ts, the pyrosequencing signal received from base incorporation is indiscernible from what the platform expects to receive from a homopolymer of 8 consecutive bases.

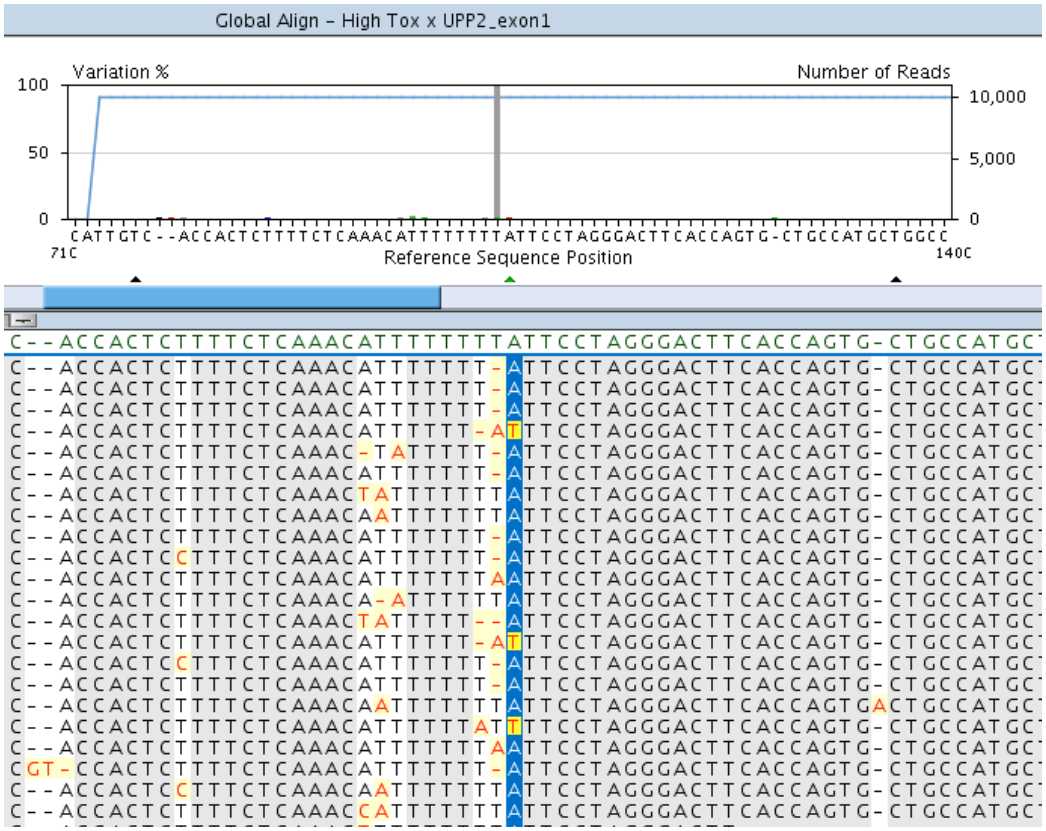


Figure 2.4: Pyrosequencing (indirectly) measures the amount of inorganic phosphate created by the incorporation of a base into a complementary strand, and this amount can be misinterpreted as the wrong number of bases being incorporated when multiple bases incorporate in a single wash (known as a "homopolymer").

The presence of any one of these criteria scored an AVA-identified variant as false. To ensure consistency, each variant was scored by two researchers independently (myself and Dr Palles), and re-checked together when discrepancies existed.

2.4 Imputation and Association Testing of Imputed Data

Imputation of genotypes with IMPUTE2 is employed in the first three results chapter of this thesis. This section briefly summarises this method, while the analyses performed to determine the specific imputation and association testing strategy are discussed in Chapter 4.

Imputation Technology

Imputation is an *in silico* method that estimates the genotype of some of a patient's non-genotyped loci using three sources of information: other genotyped loci of the patient, partial genotype information of other patients within the GWAS and dense genotype information for a reference set of samples (typically sequenced patients) (93; 133). The imputation method is useful because it can fill in missing genotypes for the purpose of fine mapping regions of interest or increase sample size at a locus for facilitating meta-analysis when studies have performed genotyping of different variants (ie, because they used different genotyping platforms) (92). In the case of data for QUASAR2, sample imputation was particularly useful as three different genotyping arrays were used across the sample set.

The full imputation method uses a patient's known genotypes in a particular genomic region to assign the two known (sequenced or densely genotyped) haplotypes that are most likely to have been inherited by the patient, including the sequence information within these haplotypes for missing loci. Historically, imputation was performed in a single step, in which genotype data from a patient were matched to reference haplotypes from which genotypes were filled in. More recently, imputation is split into a pre-phasing and imputation step. During pre-phasing, a patient's SNP array data are used to estimate haplotypes from within the set of GWAS patients.

This requires a minimum of 100 samples as the information from the dataset as a whole is used in the process. The genomic regions used are relatively small, and thus a patient will be assigned a "mosaic of haplotypes" for each chromosome during phasing. The outputs of this computationally intense step (ie, the established haplotypes) then exist and can be used to impute missing genotypes from reference panels such as the 1000 Genomes Project (91). Pre-phasing allows researchers to update their imputation datasets to increasingly more accurate and dense panels with minimised computational analysis; computational efficiency improves by an order of magnitude in exchange for minimal changes to accuracy (91).

Imputation Process and Association Testing of Imputed Data

For this thesis, pre-phasing of QUASAR2 patients was performed using SHAPEIT2 (48), run by Dr Palles using 940 QUASAR2 patients on one of three SNP arrays. Phasing was performed separately for patients genotyped on each of the three SNP arrays used. Imputation was performed using IMPUTE2 (91), using the August 2012 release (build 37/hg19) of the 1000 Genomes Project as a reference panel. Whole genome imputation was performed in a sliding window of 5 mega-base chunks across each chromosome. For imputation, Dr Palles and I together modified a perl script that was shared by Dr Kate Elliot at the Wellcome Centre, which Dr Palles ran. An effective population size of 20,000 was used, as was a 250kb buffer region, as suggested by the software developers. I performed all subsequent analyses, including association of imputed data using SNPTEST2 (135), using a script that Dr Palles and I jointly wrote, and which I subsequently modified for each of the dozens of association testing iterations run. Features of our method are highlighted below:

1. `awk` was used to identify imputed insertion/deletion SNPs – those that contained "D" or "I" in their name – so as to remove them (SNPTEST can only work on one variant per locus).

2. `awk` was used to identify duplicate SNPs – those with the same rsID – so as to remove them.

3. `awk` and `fgrep` were used to identify SNPs at the same position as another SNP (such as tri-allelic SNPs) – so as to remove them.

4. `awk` was used to identify SNPs with IMPUTE2 info scores below 0.95 for patients on *any* of the three arrays – so as to remove them (see the following section for further details).

5. The SNPTEST command was run using IMPUTE2 genotype probabilities and clinical data from patients on all three genotyping arrays at once, in parallel on a cluster computer for each of 593 chromosomal chunks. This analysis was looped over for each tested phenotype (eg, diarrhoea), for each tested outcome variable coding (eg, grade 012v34), for each selected genotype model (eg, additive).

6. Downstream quality control filtering was applied (see the following section for further details).

In summary, as developed in Chapter 4 and applied to both Chapter 4 and Chapter 5, for QUASAR2, imputation was performed separately for each of the three QUASAR2 SNP arrays, using the standard 5 mega-base sliding window SNP interval. Association testing by SNPTEST2 was performed by analysing imputation genotype probabilities for patients on all three arrays; an IMPUTE2 info score of 0.95 or greater (for each individual array) was used as an inclusion criteria. Following running SNPTEST2, additional downstream filtering was performed, namely excluding SNPs with MAF<0.01 and IMPUTE2 info scores <0.95 (pooled across all three arrays).

2.5 Statistical Methods

Population Measures

Hardy-Weinberg Equilibrium

Hardy-Weinberg equilibrium (HWE) states that allele frequencies in a population will remain constant over time, so long as the population is large, has random mating and mutation and is under no selection or migration pressure. The equilibrium is shown in Equation 2.3, where "p" represents the frequency of the major allele and "q" represents the frequency of the minor allele (ie, MAF), from which the expected frequency of the major homozygotes, heterozygotes and minor homozygotes can be calculated.

$$p^2 + 2pq + q^2 = 1 \tag{2.3}$$

Linkage Disequilibrium

Linkage disequilibrium (LD) is the non-random association of alleles between two loci. 1000 Genomes and HapMap3 LD information was downloaded from the Broad Institute's SNAP program (`http://www.broadinstitute.org/mpg/snap/`), either "pairwise LD" between two SNPs or "proxy search" to identify SNPs in strong LD. Two measures of LD are used, D' and r^2, shown in Equation 2.4 and 2.5. D' of 1 or -1 means no evidence exists for recombination between loci and high D' can mean that two markers are good proxies; however D' estimates are not reliable for rare variants. r^2 ranges from 0 when markers are in perfect equilibrium to 1 when they are perfect proxies for one another.

$$D'_{AB} = \begin{cases} \frac{D_{AB}}{\min(p_A p_B, p_a p_b)} & D_{AB} < 0 \\ \frac{D_{AB}}{\min(p_A p_B, p_a p_b)} & D_{AB} < 0 \end{cases} \tag{2.4}$$

$$r^2 = \frac{D_{AB}^2}{p_A\,(1-p_A)\,p_B\,(1-p_B)} \tag{2.5}$$

Heterozygosity

In quality control of GWAS data, per-patient heterozygosity can be measured according to Equation 2.6, which calculates the proportion of genotypes that are heterozygous for an individual. Whereas heterozygosity in a population might indicate inbreeding, high heterozygosity in an individual suggests that genotyping quality is poor (such as from low quality DNA) and calls are not clustering accurately. For the equation, N is the number of non-missing genotypes and O is the observed number of homozygous genotypes, each for an individual (7).

$$\text{mean heterozygosity} = \frac{N-O}{N} \tag{2.6}$$

Population Structure

Association studies identify allele frequency differences between cases and controls, including those differences that are unrelated to case/control status, such as from ethnic or genotyping platform differences – known as "population structure." Thus, prior to testing associations, population structure in cases and controls was examined by performing principal components analysis (PCA) using EIGENSTRAT (from EIGENSOFT) software (159), in which genetic data from study patients are clustered alongside controls of different ethnic origin from the HapMap project according to two dimensional vectors (PCs) that each explain the maximum amount of remaining variation among genetic datapoints. Ten PCs were observed in pairs, but primarily the first and second pairs were used to detect outliers. The full EIGENSTRAT method, which consists of adjusting genotypes according to population differences, was not used; rather, PCA plots were inspected for case/control population structure differ-

ences and outliers were removed prior to association testing. For heritability analyses, the first ten PCs were used as covariates.

Power Calculations

Power calculations were performed using in-house software developed at London School of Hygiene and Tropical Medicine by Dr Dan Altman. This method identifies the probability that an association test will reject the null hypothesis, given that the null hypothesis is in fact false. In other words, the method calculates how capable an experiment is of finding a positive association, given sample size, case/control ratio, variant allele frequency, and variant effect size. A table of calculated sample sizes needed to identify statistically significant associations with various toxicities with 90% power is provided in Appendix.

Model Selection

The goal of a statistical model is to accurately, efficiently and interpretably describe the relationship between your independent data (genetic data) and dependent data (outcome variable).

Covariates

Spurious associations can occur when both the outcome variable and genetic data are associated with an extraneous variable. Such confounding can be minimised by appropriate covariate use, in which a multivariate model recognises that the association between a test allele and outcome variable is due to a shared association with a third variable. However, overfitting a model with too many parameters (ie, covariates) can reduce power (154), such as by providing an overly-precise description of the data that can not be generalised to the greater population. To balance this trade-off, this thesis applies Akaike information criterion (AIC), which measures the

probability that included variables predict the outcome, penalised by the number of variables included in the model. Equation 2.7 provides the AIC equation, in which L is the likelihood function and k is the number of parameters in the model; the preferred model is that which minimises AIC (4). AIC risks selecting too large a model and thus Bayesian information criterion (BIC), which more stringently adjusts for the number of parameters, is sometimes used as a check. Stepwise model building was performed in R, either manually or with `fastbw` of the `rms` package, minimising either p-value or AIC.

$$\mathrm{AIC} = -2\ln(L) + 2k \tag{2.7}$$

Covariate selection for the Cox proportional hazards survival model was performed in two stages: first individually significant clinical variables were determined by univariate testing in a Cox proportional hazards model, then individually significant covariates were tested jointly among other clinical variables and those that became non-significant were removed. This process was repeated for both OS and DFS. Subsequently, using scaled Schoenfeld residuals (172), non-proportional hazards was tested for each covariate. All survival covariate selection was performed in R, using the `coxph` and `cox.zph` functions of the `survival` package.

Genetic Model

The numeric coding of allele dosages affects the relationship between genotype and the outcome variable that is tested. The various genetic models used in this thesis are shown in table 2.7.

Outcome Variable Bins

Not all phenotypes follow strict case/control status. Rather, binary case/control status of some phenotypes sometimes requires binning of a quantitative variable,

Model	Ref. vs Test	Weightings
Allelic	A vs B	0 vs 1
Genotypic	AA vs AB vs BB	0 vs 1 vs 2
Dominant	AA vs AB or BB	0 vs 1
Recessive	AA or AB vs BB	0 vs 1

Table 2.7: Four genetic models are commonly used. The allelic model compares cases and controls based on the number of reference (A) and test (B) alleles in each group. The genotypic model, also known as the "additive" model, is similar, but is on a per-patient (genotype) basis rather than a per-allele basis. The dominant model tests whether any test-allele dosage has an effect on the outcome variable. The recessive model tests whether homozygote test-allele dosage has an effect on the outcome variable. This last model is often underpowered for rare variants.

which can be somewhat biologically arbitrary. For the study of toxicity, which refines a spectrum of phenotypes into discrete grades of 0, 1, 2, 3, or 4, either binary variables were formed based on clinically actionable vs non-clinically actionable levels (eg, 012 vs 34) or the phenotypes were tested as quantitative variables (eg, 01 vs 2 vs 34). Sometimes, the decision to bin variables was based on statistical power, namely that a substantial number of patients (eg, 100) existed in each comparison group, in order to avoid spurious associations.

Effect Size Estimates

Regression

Effect sizes were calculated using either regression models or as odds from a 2-by-2 table of counts. Regression was applied when using raw data to identify the effect of allele dosage on an outcome, either alone or adjusted for covariates. Given our aim of testing the effect of allele dosage (usually binary or quantitative) on case/control status (binary) or toxicity grade (quantitative), logistic regression was preferred to linear regression. Logistic regression follows Equation 2.8, where the lefthand side of this equation, known as the logit function, provides the logarithm of the odds. Analysis was implemented in R using the `glm` or `Glm` function or in PLINK using the

–logistic flag. Significance of individual parameters (eg, SNP genotype) was calculated by comparing a Z score (the estimate divided by its standard error) to a two tailed Z distribution.

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_1 + \beta_2 x \tag{2.8}$$

Cox Proportional Hazards

Cox proportional hazards survival analyses were performed under an additive model in which genotypes were coded as 0, 1 or 2 based on allele dosage per patient, for each SNP. Age, stage and gender were applied as covariates for the analysis of overall survival, while disease site and stage were applied as covariates for the analysis of disease free survival. Time-to-event periods were conditioned on the duration of time from a patient's date from diagnosis to at risk period. Survival analyses were performed in R using the `coxph` function of the `survival` package. Dr Jean-Baptiste Cazier of the Bioinformatics Core at Oxford's Wellcome Trust Centre and Department of Oncology helped me modify a single-SNP analysis script into a genome-wide analysis script, for creation of association test outputs.

2-by-2 Contingency Tables (From Counts)

When raw data were not available and/or no adjustment was either needed or possible for covariates, effect sizes were estimated from allelic counts stratified by case/control status, entered into 2-by-2 tables. When analysing a single study only, a simple odds ratio was calculated, which compares the odds of case status to the odds of control status for each allele group; alternatively, relative risk would describe the relative incidence of case/control status. Significance was calculated by Pearson chi-squared, calculated using Equation 2.9, which uses the difference between the observed and expected values in each cell of a 2-by-2 table. When compared to a chi-squared

distribution, it tests the null hypothesis that the odds ratio equals 1 (no effect).

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \tag{2.9}$$

Further effect size estimates are discussed in the Meta-Analysis section below.

Meta-Analysis

Fixed-Effect

When combining data from multiple studies into one overall estimate of an effect size, meta-analysis was performed, using either counts or summary statistics. When all included studies were expected to be samples from a single (identical) population with one effect size, fixed-effect meta-analyses were used to calculate the pooled odds ratio, according to the Mantel-Haenszel method of weighting according to sample size (ie, not within-study inverse variance), shown in Equation 2.10. Significance was determined by comparing a Z score (the pooled estimate divided by its standard error) to a two tailed Z distribution. 95% confidence intervals were also calculated using two tails, as the pooled estimate +/- 1.96 times the standard error. Meta-analysis was performed using the four-variable option of the `metan` command in Stata (83), in which each variable corresponded to allelic counts in high and low toxicity patients (or, in the recessive model, to rare homozygotes versus all others).

$$\mathrm{OR}_{MH} = \frac{\sum_{i=1}^{k} \left(\frac{a_i d_i}{n_i}\right)}{\sum_{i=1}^{k} \left(\frac{b_i c_i}{n_i}\right)} \tag{2.10}$$

Random-Effects

When studies were expected to be samples of similar (but non-identical) populations with similar (but non-identical) effect sizes, random-effects meta-analysis is the preferred method. For example, this method might have been used to compare the effect

of a genetic polymorphism on patient toxicity across multiple treatment regimens. In this thesis, no studies were meta-analysed that we expected *a priori* to be samples from non-identical populations. However, when evidence of inter-study heterogeneity existed (see below), random-effects estimates were calculated and compared, calculated according to the methods of DerSimonian and Laird. This method differs from the Mantel-Haenszel method in that study weightings include an added term for between-study variance, which increases the overall variance value in the model and thus widens the pooled estimate's confidence intervals. However, meta-analysis of heterogeneous results are harder to interpret and possibly invalid (5; 119), and thus random-effects results were used with caution.

Heterogeneity

Heterogeneity was assessed using I-squared and the heterogeneity p-value. I^2 identifies the proportion of total variation among study estimates that is due to inter-study heterogeneity (ie, not chance) (86). In other words, it is the proportion of the variance that results from inter-study differences in effect size estimates, though not individual within-study variance. As dispersion among effect size estimates increases – a problem, given that studies analysed by fixed effect meta-analysis are meant to be samples from a single distribution – so does I^2. I^2 is defined in Equation 2.11, in which Q is a measure of inter-study heterogeneity (ie, observed variance) (85). Q is calculated by summing the square of the difference of each study's effect estimate from the mean effect estimate multiplied by the study's weight (ie, sample size for Mantel-Haenszel method). In this thesis, meta-analysis with I^2 values of 75% were considered to be highly heterogeneous (86), though a qualitative assessment of studies was also applied.

$$I^2 = \frac{(Q - df)}{Q} \tag{2.11}$$

Publication Bias

Missing data can bias results silently (ie, without impacting confidence intervals). For meta-analysis, the greatest concern is under-representation of null results, particularly those from small studies – known as "publication bias." One means of identifying publication bias is to create funnel plots, in which the x-axis is the effect size estimate and the y-axis is the sample size (61). Without publication bias, this plot should look like an equilateral triangle (or funnel) in which the large (and assumed more precise) studies form the tip and the smaller (less precise) studies form a broad base. Publication bias is expected to exist when an acute or right triangle is formed due to an absence of small studies that report a null effect. Funnel plots were created in Stata using the `metabias` command and specifying `egger` to fit a regression line. Formally, this line is tested for funnel plot asymmetry using Egger p-values, which test the weighted regressions of effect size on standard error to see whether estimate size/direction are related to sample size (187).

Multiple Testing

GWAS

Type I error (ie, false positives) is an expanding concern in genetics as sequencing prices fall. Due to multiple testing, the conventional $p<0.05$ significance threshold applied to the results of one analysis can hardly ever be used. Rather, various multiple testing corrections exist, often in the form of a more burdensome (ie, lower) p-value for significance that must be achieved in order to trust a genetic association. The most stringent method is the Bonferroni correction, in which each test is considered as fully independent and thus significance is achieved with a p-value that is lower than 0.05 divided by the number of tests. Due to extensive linkage disequilibrium across tested genetic markers, however, the Bonferroni correction is usually too strong and

thus increases the risk of type II error (ie, false negatives). Thus in GWAS, where sometimes millions of markers are tested, the number of independent tests corrected for is capped at 1 million and thus $p<5x10^{-8}$ is used in order to be 95% sure that an association is not the result of chance (108). This threshold is applied to the GWAS results in this thesis.

Background inflation of results was visualised using Manhattan plots, in which the x-axis is each tested SNP's position (ie, chromosome number and coordinate) and the y-axis is its association test's significance (ie, $-log_{10}(P)$). Inflation was further visualised using QQ (quantile-quantile) plots in which the x-axis is the expected (chi-squared) distribution of association test p-values and the y-axis is the observed distribution of association test p-values; deviation from the x equals y line suggests inflation due to population structure or other confounders. Inflation was more formally tested using λ, according to Equation 2.12. Plots were made in R using modified scripts from Dr Stephen Turner (`www.http://gettinggeneticsdone.blogspot.co.uk/`). λ was calculated in R using Dr Yurii Aulchenko's `GenABEL` package, which applies a 1 df expected χ^2 value of 0.456.

$$\lambda = \frac{\text{median}\left(\chi^2_{obs}\right)}{\text{median}\left(\chi^2_{exp}\right)} \tag{2.12}$$

Candidate-Based Studies

When adjusting for multiple testing for candidate gene studies, this thesis takes a conservative approach, assuming that SNPs genotyped directly on an array platform are independent and thus applying Bonferroni correction based on the number of directly genotyped SNPs (ie, not imputed) within the target regions.

Known Markers

For analysis of previously studied variants, many of which had a biological candidate basis and/or a previously suggested association, a false discovery rate (FDR) was applied. This method calculates a sliding p-value significance threshold according to Equation 2.13 (17). Here, the p-values for all tests are ordered from least to most significant and a p-value threshold of c_i is applied to each. Associations with p-values that fall below this threshold are considered true. In other words, the p-value for the least significant test must be 0.05 to be considered significant and the p-value for the tenth most significant test must be 0.005 to be considered real. In modified versions, a value "q" is applied, which is the proportion of discoveries one allows to be false. FDR was implemented using the `smileplot` command in Stata and only considering as real those associations that fell below the calculated threshold. For the calculation of this threshold in the meta-analysis, the distribution of p-values was formed using results from the "global" toxicity analyses. This measure was applied to all toxicity phenotypes, given their non-independent nature with the global measure.

$$E(Q) \leq \frac{m_0}{m} q^* \leq q^* \tag{2.13}$$

Heritability

SNPs with genome-wide significant associations often in aggregate explain only a fraction of the heritability (ie, the proportion of phenotypic variation that results from genetic variation) thought to exist based on twin studies, with a typical locus explaining about 0.25% of a disease's phenotypic variance (182). It is expected that some of this "missing heritability" remains within the moderately common SNP level, but that statistical power prevents identification and validation of SNPs with small

effect sizes (eg, OR=1.05). For many studies, such as drug toxicity or cancer survival, family-based studies are unrealistic. An alternative means of predicting heritability is to use genome-wide SNP data to calculate genetic similarity (ie, "relatedness" or "kinship") among samples and then to compare this relatedness to the amount of phenotypic similarity observed among samples.

Weighted Kinship Matrix

Prior to calculating heritability, it is advisable to define the linkage disequilibrium that exists among SNPs and to use this to weight kinship. If LD remains undefined, bias can be produced in which the relative contribution to heritability estimates of causal variants in areas of strong LD are over-estimated and that of causal variants in regions of weak LD are under-estimated; this is because tagging SNPs are similarly observed to associate with phenotype, but the total genetic component of heritability is fixed at the variance of the distribution of alleles in the population (183). This bias can be corrected, however, using linkage disequilibrium adjusted kinships (ie, LDAK), which weights SNPs according to how well they are tagged by neighboring genotyped SNPs (183). For this thesis, LDAK was performed according to the methods and software of Dr Doug Speed (`http://dougspeed.com/ldak/`), which calculates SNP weightings and kinships as inputs for SNP heritability analysis.

SNP Heritability Estimation

Heritability was then measured using a restricted maximum likelihood (REML) procedure that (a) first fits the effects of all SNPs simultaneously in a single mixed linear model that incorporates both covariates (eg, sex, age and principal components) and how each individual's genotype for each SNP differs from the average genotype, in order to explain the phenotypic variation observed, then (b) uses the maximum likelihood approach to estimate values for each variable (218). In max-

imising this likelihood, genetic heritability is estimated. For this thesis, REML was performed according to the methods and software of Dr Jian Yang and Dr Peter Visscher, known as genome-wide complex trait analysis (GCTA) (`http://www.complextraitgenomics.com/software/gcta/reml.html`). For analysis of toxicity, both binary grade 012v34 toxicity status was used (to match clinical relevance) and quantitative grade 0v1v2v3v4 status was used (as is far less prone to bias). The GCTA software is not designed for time-to-event data (which requires each of a time and event/censoring variable), and as such survival heritability estimates were attempted using binary six-year OS and DFS data.

2.6 Results Interpretation Tools

In addition to the various analysis softwares used to perform statistical analyses (eg, PLINK, Stata and R), this thesis used multiple online tools for the interpretation of results. These tools are employed repeatedly throughout this thesis and are summarised in Table 2.8.

Resource	Applications	Example of Use
UCSC Bioinformatics	Genome Browser	Visualise sequence, SNPs, ENCODE tracks, and more
http://genome.ucsc.edu/	Table Browser	Obtain DNA sequence and annotation
	Custom Tracks	Align sequence data to genome, to visualise coverage
	Lift Over	Convert between genomic builds (eg, hg18 and hg19)
	In Silico PCR	Visualise primer start/end coordinates relative to gene
	Blat	Identify reference data that matches DNA sequence data
ANNOVAR (summarize_annovar.pl function)	SIFT	Predict affect of amino acid substitutions (SIFT and Polyphen) and indels
http://www.openbioinformatics.org/annovar/	Polyphen	(Mutation Taster) on protein function, based on conservation, structure,
	Mutation Taster	function and biomedical databases
NCBI	dbSNP	Allele frequency and annotation for SNPs based on rsID
http://www.ncbi.nlm.nih.gov	dbSNP rslist	Identify chromosomal coordinates based on rsIDs
	RefSeq	Download reference DNA/protein sequence accession numbers
	Blast(x)	Identify reference data that matches DNA (protein) sequence data
	COBALT	Multiple sequence alignment
ClustalW	ClustalW	Multiple sequence alignment (COBALT alternative)
http://www.ebi.ac.uk/Tools/msa/clustalw2/		
Broad Institute	SNAP	Identify LD between enetered SNPs or search for SNPs in LD with entered SNP
http://www.broadinstitute.org/	Haploview	Visualise regional LD plots based on 1000 genomes or entered data
UMich LocusZoom	LocusZoom	Create regional manhattan plots annotated with LD and recombination sites
http://csg.sph.umich.edu/locuszoom/		

Table 2.8: Various online results interpretation tools were used in order to provide a tangible sense for identified variants and prioritise identified variants for follow-up sequencing.

Chapter 3

Systematic Review and Meta-Analysis of Genetic Markers of 5-FU Toxicity and Investigation in QUASAR2

3.1 Introduction, Aims and Overview

Despite a decade of publications on germline genetic markers of 5-fluorouracil (5-FU) toxicity, little is known convincingly. As shown in Introduction (Table 1.5), while 16 polymorphisms have been reported on for an association at $p<0.10$ and/or are included in commercially available kits for the prediction of 5-FU-related toxicity, none have reached statistical significance when adjusting for even one test each. Further, the reported methods of study – including the particular phenotypes analysed, the grouping of toxicity phenotypes, the grouping of toxicity grades, and the genetic model applied – have been inconsistent, suggesting further testing exists than is readily apparent and limiting direct replication even when variants have been analysed in

multiple studies. As a result, by merely reviewing the existing body of literature, it is difficult to identify which polymorphisms are true predictors of toxicity. In addition, while some studies perform pharmacogenetic analyses across multiple 5-FU monotherapy and/or combination therapy regimens, no report has justified the approach and it is unclear whether the results from one 5-FU-based regimen are generalisable to others.

Using a systematic review and subsequent meta-analysis, this chapter sets out to answer key questions regarding both the effect that putative toxicity variants have and the appropriate methods for the study of 5-FU toxicity biomarkers. Namely, the aims of this chapter are to:

1. Identify all previously studied 5-FU toxicity variants.
2. Analyse these variants in the QUASAR2 dataset of capecitabine patients.
3. Collect data for these variants from all previous investigations of each variant.
4. Meta-analyse these variants in capecitabine patients (our primary analysis).
5. Meta-analyse the results of other/all 5-FU-based regimens, as appropriate.
6. Identify whether data can be further combined across 5-FU-based regimens.
7. Identify whether a particular toxicity phenotype drives the effect observed in a global measure of all toxicity phenotypes.
8. Measure the ability of the identified markers to predict toxicity as a diagnostic panel.

In order to achieve these aims, this chapter aggregates data from 4,855 patients in QUASAR2 and 16 published studies in which patients received either a 5-FU monotherapy regimen (capecitabine, infusional or bolus) or combination regimen including (only) one additional cytotoxic agent (oxaliplatin or irinotecan). Toxicity

biomarker results for QUASAR2 have never been reported previously; as in the rest of this thesis, QUASAR2 patients given capecitabine monotherapy or combination therapy with bevacizumab were treated as a single group, given that the toxicity profiles of these drugs overlap minimally and only 5-FU-related phenotypes are analysed. These aggregated patient data are then used to identify the scope of the meta-analysis, namely that 36 variants have been previously studied for an association with 5-FU toxicity in at least one study and that weak evidence exists for the ability to pool data from patients on different 5-FU regimens in the analysis of these variants. Results for each of these 36 variants are then analysed in the QUASAR2 dataset and by separate meta-analysis for each of the collected 5-FU-based regimens, for both a measure of "global 5-FU toxicity" and for each of the common individual toxicity phenotypes (eg, diarrhoea). Multiple testing is controlled for by calculating a significance threshold based on false discovery rate. Ability to further share data across regimens is then assessed. Finally, the markers with strong evidence among capecitabine patients are modeled for predictive sensitivity/specificity and area-under-curve in the QUASAR2 dataset, applying the effect estimates from the second largest capecitabine patient set. By doing so, this chapter aims to draw a line in the sand regarding the field as it stands given both current publications and data on previously studied markers from the QUASAR2 trial. PRISMA guidelines were followed.

3.2 Challenges Particular to this Analysis

Generally for a meta-analysis, one outcome is tested and much of these data are available from publications and supplementary information. This analysis differed from previously published reports, however, in that multiple outcomes (ie, an undefined number of previously tested polymorphisms) were being analysed. Further, as the nature of genetic analysis and reporting is that manuscripts focus on the most im-

portant findings among many variants being tested, it was obvious to us that many data existed beyond the reaches of openly available data. An additional challenge for this analysis was that published analyses used inconsistent outcome variables and statistical models – five toxicity grades would be collected (grade 0 to 4) for as many as seven 5-FU-related phenotypes, and sometimes non-5-FU-related phenotypes in addition, then binned according to each author's method; genetic variants could be previously analysed in any of an allelic, genotypic, dominant, or recessive model. Thus it was often necessary for authors to recode their data or analyses such that we could compare similar measures across studies.

In order to obtain the necessary data for the meta-analysis, the data aggregation process had the following steps and goals:

1. Perform a literature search in order to:
    - define the list of previously studied variants for 5-FU toxicity
    - identify all relevant studies for each variant
2. Contact authors of each study in order to:
    - form a collaborative group
    - identify whether additional variants were studied beyond those reported on
    - collect data from all studied variants
    - standardise the coding of all collected data

3.3 Literature Search

We sough to identify 5-FU toxicity variants in English language papers published up to and including 2012. The literature search used for this systematic review is reproduced in full in Appendix. This literature search included 139 separate searches

in total. First, in PubMed, combinations of the search terms "toxicity" and "5-fluorouracil", "5-FU" or "capecitabine" were searched (3 searches). From the collected studies, a list of 8 previously-studied *genes* was formed. The initial search terms were then re-searched in combination with each of these gene names or commonly used gene symbols (24 searches). From all collected studies, a list of 56 previously studied *polymorphisms* was formed. The search terms from the initial search and second search were then re-searched for each of the identified polymorphisms (eg, "toxicity and 5-FU and CDA or CDD and 943insC or rs3215400"; 56 searches). This formed a list of all previous *studies* for each polymorphism. Finally, in order to complete this list of studies, the first literature mention of each polymorphism was identified and the "cited by" option of GoogleScholar was used to list all subsequent studies that cited this initial study (56 searches), and then the references list of each collected study was reviewed for further relevant citations.

3.4 Study Filtering and Formation of Patient Sets

Not all identified studies and polymorphisms were included in the meta-analysis. Rather, stringent inclusion/exclusion criteria were applied that would prevent bias in the final analysis. Figure 3.1 provides a schematic of the study and polymorphism filtering process.

First, studies were considered for inclusion if they:

1. included patients on a 5-FU-based regimen with at most one additional cytotoxic agent

2. analysed genetic polymorphisms in relation to these regimens

3. performed analyses using one or more of the six common 5-FU-related toxicities as an outcome variable:

	Gene symbols	Studies	Study Filters			Patient Sets	Patient Set Filters	
Drug Activation	CES2	2	all studies pass	2	After citation collection and study filtering, 28 studies remained, many listed for multiple genes. (These studies in aggregate investigated 59 variants for an association with 5-FU toxicity, as shown in later tables/figures.) Remaining studies were then split into separate patient sets for each 5-FU treatment regimen used within the study, then subsequently filtered for exclusion criteria. Figure now includes QUASAR2 in patient set counts.	3	all study arms pass	3
	CDA (CDD)	2	all studies pass	2		3	all study arms pass	3
	UMPS (OPRT)	4	not CEU = 2	2		5	overlap = 1 no data provided = 2	2
	TYMP (TP)	1	all studies pass	1		2	all study arms pass	2
Drug Target	TYMS (TS)	27	n<30 = 3 not CEU = 2 unpublished = 2	20		32	>2 treatments = 1 overlap = 1 n<30 = 1 no data provided = 8	21
	MTHFR	21	n<30 = 1 not CEU = 1 unpublished = 1	18		28	>2 treatments = 1 no data provided = 7 overlap = 2	18
Drug Catabolism	DPYD (DPD)	34	n<30 = 7 not CEU = 2 unpublished = 2 opportunistic = 9	14		28	>2 treatments = 1 no data provided = 4 mixed regimens = 1 n<30 once split = 8 overlap = 1	13
	DPYS (DHP)	2	n<30 = 1 opportunistic = 1	0		0	No study arms remain	
Studies/patient sets: (unique)		49	→ *Study Filtering*	28	→ *Studies split into patient sets of a single regimen*	47	→ *Patient Set Filtering*	24

Figure 3.1: The study filtering process for the meta-analysis is shown. Of the 49 identified publications, 28 remained after applying per-study quality control filters. Subsequently, in order to separately study patients on different 5-FU-based regimens, these studies were split into 47 patient sets, each comprised of patients on the same 5-FU-based regimen within a study. After applying further quality control filters on a per-set basis, 24 patient sets remained for analysis, including QUASAR2.

- diarrhoea, nausea/vomiting, mucositis/stomatitis, neutropaenia, thrombocytopaenia, and hand-foot syndrome

49 publications met these initial inclusion criteria, which identified relevant polymorphisms in 8 genes. Studies were then restricted to those with:

1. a sample size of at least 30 patients
2. Caucasian patients
3. prospective study design
4. publication in a peer-reviewed journal (ie, not a thesis or conference abstract)

28 publications met these initial inclusion criteria, in aggregate analysing 59 polymorphisms in 7 genes (listed in the next section) (2; 3; 19; 20; 25; 30; 31; 37; 39; 47; 50; 62; 71; 72; 77; 79; 115; 116; 138; 139; 140; 145; 162; 166; 168; 173; 175; 220). Some publications included patients given one of multiple 5-FU-based regimens. In order to analyse patients on each 5-FU-based regimen separately, studies were then divided into "patient sets," with each including patients on one 5-FU-based regimen. At this point, the QUASAR2 patients were also added as a capecitabine patient set. Genotyping for QUASAR2 is described elsewhere in this thesis; genotype data for QUASAR2 were either extracted from the Illumina SNP arrays, imputed using the August 2009 release of 1000 genomes, genotyped using KASPar, or genotyped by previously reported methods, as shown in Appendix. I genotyped two of these variants by KASPar and the rest were available from previous work. Together with the 28 publications thus far included, 47 patient sets were formed.

Further filtering was then applied to the patient sets, such that each set:

1. contained at least 30 patients
2. had no patients overlapping with another set (in the case of overlap, the larger patient set was selected)

3. provided at most one cytotoxic agent additional to 5-FU

4. reported or provided an unadjusted odds ratio and a measure of variance, or genotype counts by toxicity type

Including QUASAR2, 24 patient sets (from 17 studies) were included in the meta-analysis, based on these inclusion criteria and availability of data from publications or contact with authors. An overview of these patient sets is provided in Table 3.1. In summary, global grade 3+ toxicity was collected for 20 patient sets, comprised mostly of CRC patients, most of whom had metastatic CRC. Three studies included patients with breast cancer (44% to 100%) and one study included patients with multiple cancers; these four studies were included upon review of their 5-FU-based regimens for similarity to that for CRC patients.

3.5 Identification of Included Variants and Aggregation of Included Data

Authors were then contacted by email in order to identify additional studied variants beyond those made clear in the literature; in obtaining the data and preparing the analysis, correspondence with authors required about 650 emails. Data were requested either as summary statistics, counts or in raw per-patient format. When the data collection process was complete, 59 putative 5-FU toxicity variants had been identified among 28 studies, as shown in Table 3.2 and 3.3. Of these, 36 were considered "previously studied," while 19 were considered "unstudied" because they had been invariant in all previous analyses and 4 were discarded because they could not be mapped to the genome unambiguously. Of these 36 studied variants, no data could be obtained for 3 and one was filtered in its only contributing study due to Hardy-Weinberg equilibrium ($p<0.01$). This left 32 variants for inclusion in the final analysis,

Study and Treatment Categorization	Study design	Location	Patients genotyped	Males/ Females	Age mean or median* (range)	Treatment: 1st line (n)	Treatment: 2nd /3rd line (n)	Performance Status 0-2 (n, of total)	Disease	Treatments (excluded regimens not listed; full regimen description provided in Supplementary Item 5)	Non-published data provided by authors; notes†
Afzal, 2011 Validation BOL† **	Prospective (validation) cohort	Denmark	335	171/ 169	61* (20–85)	NS	NS	NS	CRC, stage II-IV	Mayo regimen (n=340)	Counts for Danish cohort (all data included in meta-analyses); Italian cohort excluded due to overlap with Gusella et al.
Boige, 2010 INF† Boige, 2010 INF+OX†	RCT (FFCD)	France	168 176	214/ 132	67* (37-80) 68* (34-83)	346	0	346	mCRC	LV5FU2 (168), FOLFOX6 (176)†	Counts "at any cycle" for all included phenotype/polymorphism combinations.
Braun, 2009 INF† Braun, 2009 INF+IR† Braun, 2009 INF+OX†	RCT (FOCUS)	59 UK centres and 1 Cypriot centre	688 175 172	691/ 325	64* (27-85)	1016	0	1016	mCRC	Arm A: FU and second line Irinotecan (n=184), Arm B FU and second line IrFu or OxFu, Arm C: first line IrFu (n=264) or OxFu (n=280)†	Raw toxicity/genotype data, separately for FU, OxFu, and IrFu regimens.
Caronia, 2011 CAP	Retrospective study of RCT	Spain	130	18/ 112	63* (28–88)	65	13	NS	mCRC (28%); breast (72%)	Capecitabine, standard (n=104) or continuous (n=26)	Raw toxicity/genotype data for all included phenotype/polymorphism combinations.
Chua, 2009 INF+OX	Three successive and prospective phase II trials	Australia	118	80/ 38	61 (31–75)	35	83	102	mCRC	FOLFOX (n=119)	
Etienne-Grimaldi, 2010 INF+OX	RCT (OPTIMOX 2)	France	117	65/ 52	67* (31–80)	103	14	111	mCRC	FOLFOX 7 regimen (n = 58) or 6 cycles of mFOLFOX 7 followed by LV–5FU (n=59)	Counts for all included phenotype/polymorphism combinations.
Glimelius, 2011 Outcome B BOL+IR† Glimelius, 2011 Outcome B INF+IR†	RCT (Nordic VI) with retrospective blood collection	Sweden and Norway	75 65	89/ 51	62 (47-75)	117	23	140	mCRC	FLIRI (n=75), Lv5FU2-IRI/FOLFIRI (n=65)	Raw toxicity/genotype data for all included phenotype/polymorphism combinations.
Gusella, 2009 BOL	Prospective cohort	Italy	130	84/ 46	64 (34-84)	NS	NS	130	CRC, Dukes B2 and C	Mayo regimen (n=130)	Counts for all included phenotype/polymorphism combinations.

Table 3.1: A summary of the studies (listed by patient set) included in the meta-analysis is shown, part 1 of 2. Cross denotes data obtained specifically for this systematic review and meta-analysis, including stratification of originally reported data into a separate patient set for each 5-FU-based regimen. NS=not stated. PFS=Performance Status. RCT=Randomised control trial. mCRC=metastatic colorectal cancer. *Median. **Data on Danish cohort provided by authors. Italian cohort reported in original publication overlaps with the included Gusella et al study and was omitted.

Largillier, 2006 CAP	Prospective cohort	France	87	NS	61 (33-84)	29	76	95	Breast cancer (100%), advanced	Capecitabine (n=105)	Counts "at any cycle" for all included phenotype/polymorphism combinations.
Lecomte, 2004 INF	Prospective cohort	France	47	47/ 43	65 (42–83)	81	9	NS	Stage II/III CRC (43%) and mCRC	5-FU continuous (n=2), LV5FU2 (n=45)†	Raw toxicity/genotype data for INF regimens alone, for all included phenotype/polymorphism combinations.
Martinez-Balibrea, 2008 CAP+OX Martinez-Balibrea, 2008 INF+OX	Randomised phase III study prospective cohort	Spain	47 49	60/ 36	63* (36–82)	70	26	96	mCRC	XELOX (n=47), FUOX (n=49)	Raw toxicity/genotype data for all included phenotype/polymorphism combinations.
Martinez-Balibrea, 2010 INF+IR	Randomised phase III study prospective cohort	Spain	149	90/ 58	61 (28-75)	149	0	NS	mCRC	FUIRI (n=77), FOLFIRI (n=72)†	Raw toxicity/genotype data for all included phenotype/polymorphism combinations.
McLeod, 2010 BOL+IR McLeod, 2010 INF+OX	RCT (N9741)	North America	103 248	406/ 214	61 (26-85)	442	78	520	mCRC	IFL (n=114), FOLFOX (n=299)†	Counts separately for BOL+IR and INF+OX arms, for all included phenotype/polymorphism combinations. Toxicity incidence data by treatment arm.
QUASAR2, 2013 CAP	RCT (QUASAR2)	UK	927	531/ 396	64 (27-85)	927	0	927	CRC, stage II/III	Capecitabine (n=436), Capecitabine + bevacizumab (n=491)	Raw toxicity/genotype data for all included phenotype/polymorphism combinations.
Ribelles, 2008 CAP	Prospective cohort	Spain	123	35/ 101	66* (32-90)	42	94	113	mCRC or breast (44%)	Capecitabine (n=123)	Counts for all included phenotype/polymorphism combinations.
Schwab, 2008 BOL† Schwab, 2008 INF†	Prospective cohort	Germany	338 333	383/ 300	64 (25-89)	NS	NS	NS	CRC (91%), other GI, breast (2%), and unknown primary	Weekly infusional FU and FA, in various dosages (n=265+54), daily infusion of FU and FA (n=10), daily continuous infusion of 5-FU (n=4), Mayo regimen (n=254), other bolus regimens (n=89)†	Raw toxicity/genotype data for all included phenotype/polymorphism combinations.
Sharma, 2008 CAP	Prospective cohort	Australia	55	35/ 19	72* (42-86)	39	15	51	mCRC	Capecitabine (n=55)	Counts for all included phenotype/polymorphism combinations.

spanning the 24 patient sets previously described. Some variants included in the final analysis were merely single-study reports of data, standardised to the reporting method of the meta-analysis.

Data from these patient sets were then manually extracted independently by Dr Palles and myself, one from the original publication, 10 from investigator-supplied counts and 13 from investigator-supplied individual patient data. I compared these extractions and differences were resolved in discussion.

3.6 Toxicity Incidence by Regimen and Formation of a "Global" Measure

Incidence of adverse events (ie, drug toxicity) varies by 5-FU regimen. While the differences in toxicity incidence rates across regimens are generally understood by clinicians, this understanding is based largely on results from individual phase II/III trials and the ranges have been poorly summarised in the literature. In order to determine whether different 5-FU-based regimens result in a similar phenotypic profile – or whether 5-FU-based regimens fundamentally differ and require separate analysis – the first analysis performed in this Chapter was a summary of toxicity incidence by 5-FU-based regimen. In accordance with the most universally followed clinically relevant outcome variable, toxicity events are dichotomised into grade 0/1/2 versus grade 3+ CTCAE adverse events, as would be dose limiting.

Table 3.4 shows the incidence of individual toxicity phenotypes across seven 5-FU-based regimens. Data are presented for six common individual toxicity phenotypes – those with an incidence of at least 5% in one or more of the common 5-FU-based regimens. The only toxicity phenotype that is common among all seven regimens is diarrhoea, although rates of this adverse event are approximately two-fold higher in patients on combination regimens than in patients given 5-FU monotherapy. Impor-

Included Polymorphisms	Cohen 2003	**Lecomte 2004**	**Largillier 2006**	Morel 2006	Boisdron-Celle 2007	Gross 2007	Salgado 2007	Capitain 2008	**Martinez-Balibrea 2008**	**Ribelles 2008**	Ruzzo 2008	**Sharma 2008**	**Schwab 2008**	Afzal 2009	**Braun 2009**	**Chua 2009**	Derwinger 2009	**Gusella 2009**	Goekkurt 2009	**Boige 2010**	**Etienne-Grimaldi 2010**	**Martinez-Balibrea 2010**	**McLeod 2010**	Zarate 2010	**Afzal 2011**	**Caronia 2011**	Deenen 2011	**Glimelius 2011**	Studies/Variant
CES2 823 (830) C/G 5'UTR										1																1			2
CES2 intronic rs11568314																										1			1
CES2 intronic rs11568311																										1			1
CES2 6046G>A; R270H										1																			1
CES2 6320 G/A										1																			1
CES2 intronic rs2241409																										1			1
CDA -451C>T																										1			1
CDA -205C>G																										1			1
CDA 5'UTR SNP rs602950																										1			1
CDA 943insC										1																1			2
CDA*2; 79A>C; K27Q										1																1			2
CDA 575 C/T										1																			1
CDA 794 G/A										1																			1
CDA 771 C/G										1																			1
UMPS (OPRT) 638G>C																									1				1
UMPS (OPRT) 1336A>G																									1				1
TYMP intronic rs470119																										1			1
TYMP A324A																										1			1
TYMP S471L																										1			1
TYMS 5'VNTR 3R G/C SNP		1	1						1			1						1		1	1			1	1	1			10
TYMS 5'VNTR		1	1				1		1	1	1	1	1		1			1		1	1	1	1	1	1	1		1	18
TYMS 3'UTR 1494indel6b		1	1				1	1	1		1	1	1		1			1	1	1	1	1	1	1	1	1			18
MTHFR 677C>T; A222V	1		1					1			1	1	1	1	1	1	1	1	1	1	1		1	1	1			1	18
MTHFR 1298A>C; E429A			1					1			1	1	1	1				1	1	1	1		1	1	1			1	14
DPYD *9A; 85T>C; C29R				1	1	1																	1		1		1		6
DPYD 496A>G; M166V						1							1												1		1		4
DPYD 623G>A; R208Q													1																1
DPYD 1109delTA						1																							1
DPYD 1236G>A; E412E						1							1														1		3

Table 3.2: Of the 59 polymorphisms identified by the literature search, 36 have been previously studied for an association with 5-FU toxicity and were included in the analysis. These variants come from at least one of 28 published studies, of which 16 were included in the meta-analysis. Bold font studies are those included in the final analysis, within which polymorphisms represented by green cells were collected, data could not be obtained for red cells and data were not requested for non-coloured cells (usually because the study was removed by filtering). V=invariant in that study. S=sequenced but with no evidence or suggestion of a frequency difference in cases and controls (ie, no association test performed). For DPYD 2846T>A, we show the alleles on the positive strand, A being the rare allele, whereas some previous reports refer to this polymorphism as 2846A>T.

Invariant Polymorphisms	Cohen 2003	**Lecomte 2004**	**Largillier 2006**	Morel 2006	Boisdron-Celle 2007	Gross 2007	Salgado 2007	Capitain 2008	**Martinez-Balibrea 2008**	**Ribelles 2008**	Ruzzo 2008	**Sharma 2008**	**Schwab 2008**	Afzal 2009	**Braun 2009**	**Chua 2009**	Derwinger 2009	**Gusella 2009**	Goekkurt 2009	**Boige 2010**	**Etienne-Grimaldi 2010**	**Martinez-Balibrea 2010**	**McLeod 2010**	Zarate 2010	**Afzal 2011**	**Caronia 2011**	Deenen 2011	**Glimelius 2011**	Studies/Variant
CES2 5841 G/A										V																			0
CES2 6174 G/A										V																			0
CES2 854G/C										V																			0
DPYD 1003G>T; V335L				V																									0
DPYD 1039-1042 del TG				V																									0
DPYD 1156G>T; E386Ter				V	V																								0
DPYD 1475C>T; S492L				V																									0
DPYD 1897delC				V																									0
DPYD 257C>T; P86L				V																									0
DPYD 2657G>A; R886H				V	V																								0
DPYD 2933A>G; H978R				V																									0
DPYD 295-298del TCAT				V	V																								0
DPYD 2983G>T; V995F				V	V																								0
DPYD 601A>C; S201R				V																									0
DPYD 61C>T; R21X				V																									0
DPYD 62G>A; R21Q				V																									0
DPYD 632A>G; Y211C				V																									0
DPYD 703C>T; R235W				V																									0
DPYD 812delT				V																									0
Ambiguously Mapped																													
MTHFR 1793G>A; R549Q																							1						1
DPYD -1590 T>C				1	1																								2
DPYD IVS10-15T>C						1																							1
DPYD IVS5+18G>A						1																							1

Table 3.3: Of the 59 polymorphisms identified by the literature search, 19 were not truly studied as a marker of 5-FU toxicity, such as when a sequencing study identified the presence of a variant but did not find an allele frequency difference between cases and controls. A further 4 variants could not be mapped unambiguously to the genome. These variants come from at least one of 28 published studies, of which 16 were included in the meta-analysis. Bold font studies are those included in the final analysis, within which polymorphisms represented by green cells were collected (and in this case later discarded) and data were not requested for non-coloured cells. V=invariant in that study.

tantly, this suggests that drugs that are combined with 5-FU, such as oxaliplatin and irinotecan, might act as confounders of 5-FU-related drug toxicity, particularly as oxaliplatin induces a 5-10% incidence of diarrhoea and nausea/vomiting as monotherapy (143) and irinotecan monotherapy causes a 15-20% incidence of these same toxicities (43). As an additional confounder, it is well recognized that oxaliplatin increases incidence of thrombocytopenia to approximately 10% when used in combination with 5-FU, even though incidence of thrombocytopenia is below 5% among patients given oxaliplatin or 5-FU monotherapy (143).

In order to provide of a clinically useful understanding of toxicity genetic biomarkers, and to increase statistical power by increasing event rates, incidence of these individual toxicity phenotypes were combined for this analysis into a single measure of "global" toxicity that includes any CTCAE grade 3+ 5-FU-related toxicity at any cycle of treatment. Table 3.5 shows that the rates of dose limiting toxicity vary far less among regimens than do the underlying specific toxicity phenotypes. Contrary to the initial head-to-head comparisons of 5-FU regimens found in the literature (202), infusional 5-FU showed the lowest incidence of dose limiting toxicity, as is currently accepted clinically.

Despite interventions, onset of toxicity can be rapid and severe, resulting in approximately 0.5-1% drug-induced mortality (ie, grade 5) for monotherapy regimens, and up to 3% drug-induced mortality for combination regimens (169). Among patients included in the genetic analysis of QUASAR2, two patients died from 5-FU related toxicity (0.2%), one from respiratory failure second to neutropaenia and the other from neutropaenic colitis and left ventricular hypertrophy. Notably, drug-induced mortality does not tend to result from a single given toxicity phenotype, but rather death results from a "cluster" of toxicities occurring simultaneously or in rapid succession (165). This supports the use of a "global" measure of toxicity that captures any severe toxicity event. Further, it supports the study of predictive ge-

	Regimen	Average incidence	Studies	Patients	Incidence range (%)
Diarrhoea	BOL	12%	3	803	9-16
	BOL+IR	16%	2	178	5-24
	CAP	10%	6	1321	5-13
	CAP+OX	15%	1	47	NA
	INF	5%	5	1266	3-11
	INF+IR	16%	3	381	3-30
	INF+OX	8%	6	871	3-18
Mucositis/ Stomatitis	BOL	15%	3	803	13-28
	BOL+IR	1%	1	103	1-1
	CAP	2%	5	1234	0-5
	CAP+OX	0%	1	47	NA
	INF	2%	4	1094	1-3
	INF+IR	3%	2	315	1-4
	INF+OX	1%	4	577	0-4
Neutropaenia	BOL	7%	3	803	6-8
	BOL+IR	26%	2	178	4-42
	CAP	2%	4	1111	0-3
	CAP+OX	11%	1	47	NA
	INF	4%	4	1194	1-5
	INF+IR	14%	3	382	3-21
	INF+OX	34%	6	874	8-62
Thrombocytopaenia	BOL	0%	2	468	0-0
	BOL+IR	2%	1	103	NA
	CAP	0%	3	981	0-0
	CAP+OX	2%	1	47	NA
	INF	0%	3	1029	0-0
	INF+IR	0%	2	316	0-1
	INF+OX	5%	5	753	1-10
Nausea/Vomiting	BOL	3%	3	803	0-4
	BOL+IR	9%	2	178	7-11
	CAP	1%	5	1234	0-2
	CAP+OX	11%	1	47	NA
	INF	3%	4	1103	0-4
	INF+IR	3%	3	381	2-4
	INF+OX	6%	4	579	4-8
HFS	BOL	0%	2	468	0-0
	BOL+IR	0%	1	103	NA
	CAP	22%	6	1321	9-32
	CAP+OX	No data			NA
	INF	1%	4	1090	0-2
	INF+IR	0%	1	162	NA
	INF+OX	0%	3	526	0-1

Table 3.4: The incidence of specific toxicity phenotypes that result from treatment with various 5-FU regimens is shown, using data aggregated as part of the systematic review. Consistent with the clinical understanding and use of these 5-FU regimens, incidence of particular toxicity phenotypes varies greatly by regimen. White cells of the table show common toxicity phenotypes for each regimen, while shaded cells show the rare phenotypes, which were not considered to have sufficient power for genetic association study.

	Schedule	Average incidence	Studies	Patients	Incidence range (%)
Global	BOL	29%	3	803	21-35
	BOL+IR	No data			NA
	CAP	33%	6	1314	15-42
	CAP+OX	No data			NA
	INF	12%	5	1283	10-17
	INF+IR	30%	3	382	11-48
	INF+OX	40%	5	756	22-73

Table 3.5: The incidence of "global" toxicity, comprised of any CTCAE grade 3+ 5-FU-related toxicity phenotype at any treatment cycle, varies less by regimen than do the underlying individual toxicity phenotypes, though infusional 5-FU provides markedly lower incidence of dose limiting toxicity. Importantly, these results do not consider the treatment benefit (ie, survival) provided by each regimen.

netic biomarkers, as treatment-induced death can not be fully avoided merely by dose reducing patients who suffer early moderate toxicity events.

3.7 Assessment of Publication Bias

Publication bias was assessed for each variant using funnel plots and by calculating an Egger p-value, as described in Methods. Given that 32 variants were tested, we would expect that some would fall below the normal p=0.05 significance threshold for showing evidence of publication bias, though a strict Bonferroni correction to p=0.0016 might be too lenient. The lowest p-values obtained were for *DPYD* 1236G>A (p=0.012), *CDA* -451C>T (p=0.032) and the *TYMS* 5'VNTR repeat by counts (p=0.039), giving only minimal reason for concern about bias in the collection of data, particularly among variants with data from many patient sets. Funnel plots for the four variants considered significant in the meta-analysis are shown in Figure 3.2, in which we observe a slight underrepresentation of *positive* reporting for the *TYMS* 5'VNTR.

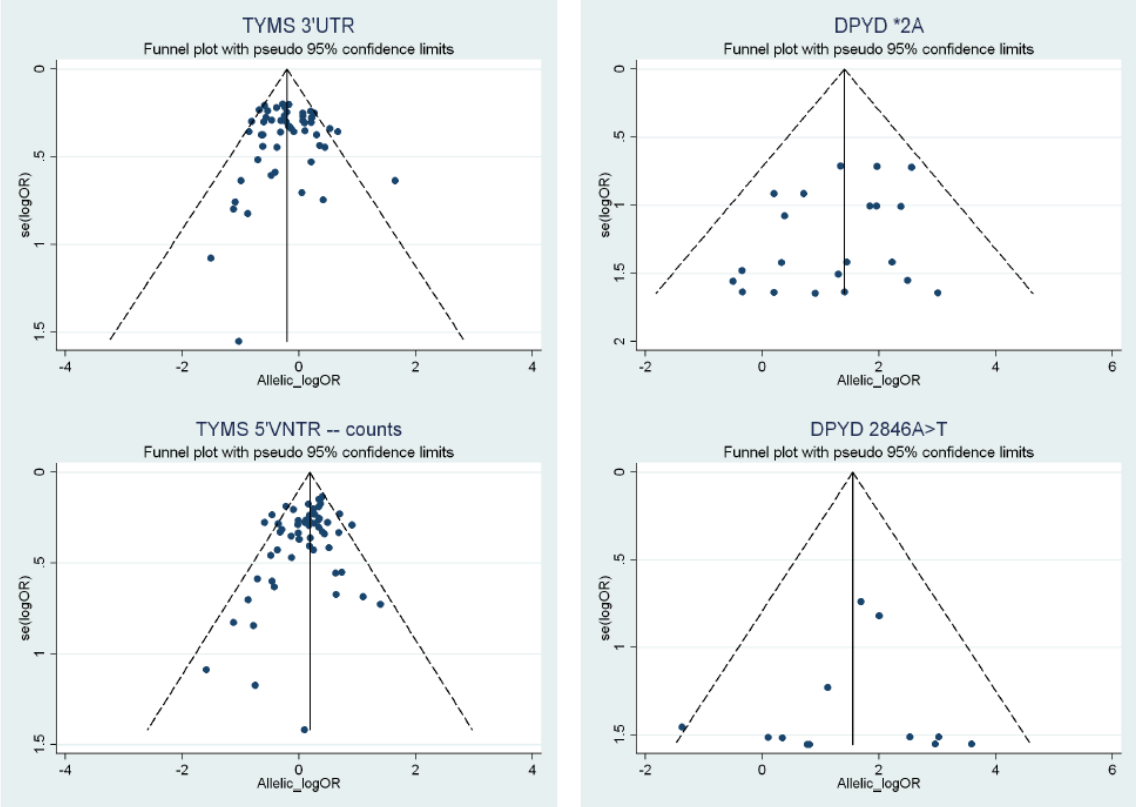


Figure 3.2: Plots and Egger p-values were produced for all variants in the meta-analysis prior to analysis in order to identify publication bias. Funnel plots are shown for the four variants identified as true predictors of 5-FU toxicity in the meta-analysis. Egger p-values among all 5-FU schedules: *TYMS* 5'VNTR=0.04; *TYMS* 3'UTR=0.76; *DPYD* *2A = 0.13; *DPYD* 2846=0.64.

3.8 Summary of Analysis Plan

Following collection of data, statistical analysis of variants was performed, first for the QUASAR2 patient set alone, then separately for each of the collected 5-FU-based regimens. Data from each patient set were tabulated as counts of the number of patients in each of the common homozygote, heterozygote and rare homozygote genotype groups, stratified by grade 0/1/2 or grade 3/4/5 toxicity, for each phenotype. I then transformed these data into allelic and recessive 2-by-2 tables per phenotype.

Where data were available for multiple patient sets of the same regimen, meta-analysis was performed. Where only one patient set was available for a given regimen/phenotype combination, lone analyses were performed according to our stan-

dardised format (the included phenotypes, phenotype pooling and toxicity grade binning). The primary analysis was to test whether each variant was associated with global capecitabine-related toxicity according to a fixed-effect meta-analysis. When evidence of high heterogeneity existed, according to I^2 or the p-value for heterogeneity, a random-effects model was also applied. Sensitivity analyses were performed, both (a) for each of the additional 5-FU-based regimens and (b) for each common (incidence >5%) individual phenotype, also in grade 012v345 format, in all regimens. Meta-analysis was performed according to the Mantel-Haenszel method, using the `metan` command in Stata. Allelic and recessive model analyses were performed, although the allelic model was our primary analysis and recessive model results were only considered where they differed substantially. Significance was determined by a false discovery rate (FDR) threshold set according to the methods of Benjamini and Hochberg (17). Using a single-study threshold of $p<0.05$ this corresponded to $p<0.0071$ for the QUASAR2 analysis, $p<0.0033$ for the capecitabine meta-analyses, and $p<0.0048$ for the non-capecitabine meta-analyses (based on analysis results). Specific statistical methods are described in Methods. Power calculations showed that we had 90% power to detect an association with global toxicity for an odds ratio of 1.5 or greater for capecitabine, for variants with a minor allele frequency (MAF) of 10%, and with MAF of 1% for an odds ratio of 3.0 or greater; sample sizes and event rates were somewhat lower for most of the other regimens (full calculation tables shown in Appendix). Results for individual 5-FU-based regimens are discussed in text using tables; forest plots of results showing the effect of polymorphisms across regimens are discussed after the individual regimen results, following a summary of significant findings.

3.9 QUASAR2 Results for Previously Studied Variants and Development of the *TYMS* Score Test

Of the 36 polymorphisms previously studied for an association with 5-FU-related toxicity, 21 were genotyped in QUASAR2 for at least 400 patients at the time of analysis. From the total patient set of 927 patients (436 capecitabine alone, 491 capecitabine + bevacizumab), 17 of these 21 variants were genotyped in at least 800 patients. In addition to global dose limiting toxicity, diarrhoea and handfoot syndrome were sufficiently common to study in capecitabine monotherapy (incidence of 10% and 22% for global grade 3+ toxicity).

Results of the QUASAR2 single-study allelic analyses for global grade 012v34 toxicity are shown in Table 3.6. Alleles of three variants were significantly associated with global grade 3+ toxicity at the $q<0.05$ ($p<0.0071$) threshold for this analysis: *TYMS* 5'VNTR 2R (OR=1.48, 95% CI 1.22-1.80, p=0.000079), *TYMS* 3'UTR 6bp ins (OR=1.67, 95% CI 1.23-2.22, p=0.00084) and *DPYD* 2846 A (OR=9.35, 95% CI 2.01-43.4, p=0.0043). The *DPYD* variant appeared stronger for diarhoea, while the two *TYMS* variants were more significant for HFS, the more common and statistically powered phenotype. None of the 18 remaining polymorphisms reached suggestive significance ($p<0.05$) for global toxicity or formal significance for individual phenotypes, though the *CES2* intronic SNP rs2241409 was suggestive for HFS and the *DPYD* *2A variant showed a large effect size. Of note, of the two patients who experienced capecitabine-induced death in the QUASAR2 trial, one had a *DPYD* *2A variant and the other had a 2846 A allele. Among the rare variants, no SNPs were formally significant under a recessive model, though *CES2* rs11568314 and rs11568311 and *TYMP* S471L showed suggestive significance for HFS, HFS and global, respectively. Of the 21 investigated SNPs, 13 have been previously reported as significant at $p<0.10$ in at least one 5-FU-based regimen for at least one phenotype, which possibly suggests

that the field as a whole has done too much testing to allow the use of a $p<0.05$ significance threshold.

Previous work has suggested that the *TYMS* 5'VNTR (173) and 3'UTR (79) polymorphisms affect levels of 5-FU-related toxicity individually, though prior to the QUASAR2 analysis, the significance levels were only modest (p=0.02 and 0.10 for global toxicity, respectively). These variants are in modest LD in QUASAR2 (r^2=0.17, D'=0.64). Further, some evidence suggests that these variants function in a haplotype, given that their individual significance is improved upon when analysed together (as low as p=0.01 for some haplotypes vs p=0.02 for the 5'VNTR and p=0.50 for the 3'UTR individually for global toxicity in the only published haplotype analysis) (116).

We therefore tested whether the *TYMS* 5'VNTR and 3'UTR variants function as a haplotype in the QUASAR2 data (results shown in table format alongside results for the same analysis in the other monotherapy regimens, in Table 3.12). First, logistic regression was used (n=474). When testing the variants individually for global 012v34 toxicity, both the 5'VNTR (OR=1.45, p=0.0070) and 3'UTR (OR=1.71, p=0.0065) were significant, as expected from the chi-squared analysis (some difference in results due to reduced sample size). When tested in a joint model, the 5'VNTR became non-significant (OR=1.24, p=0.15) and the significance of the 3'UTR was greatly reduced (OR=1.56, p=0.10). No evidence of epistasis existed. We then tested whether a trend for significance and effect was present across the four haplotypes (from least to most harmful, based on individual variant results: 3R/del, 2R/del, 3R/ins, 2R/ins) and found incremental increases in effect size (OR=1.00, 1.19, 1.53, 1.93) and modest significance (p=0.0025); however, no single haplotype was significantly associated with toxicity, the overall p-value met the FDR-adjusted threshold and the incremental increase was not observed for the individual phenotypes diearrhoea and HFS. These results together suggested that both variants might contribute to one's risk of toxicity,

Polymorphism	Toxicity	N	MAF	HWE	OR	95% CI	p-value
CDA -451C>T	GLOBAL	905	0.34	0.90	0.93	0.76-1.15	0.51
Allele 1 = T	DIARRHOEA	905	0.34	0.90	0.94	0.68-1.29	0.69
Allele 0 = C	HFS	905	0.34	0.90	0.83	0.65-1.05	0.12
CDA*2 79A>C	GLOBAL	926	0.34	0.51	0.99	0.80-1.21	0.91
Allele 1 = C	DIARRHOEA	926	0.34	0.51	1.00	0.73-1.37	0.99
Allele 0 = A	HFS	924	0.34	0.47	0.88	0.69-1.11	0.28
CES2 823 (830) C/G 5'UTR	GLOBAL	881	0.076	0.14	1.17	0.81-1.69	0.40
Allele 1 = G	DIARRHOEA	881	0.076	0.14	0.97	0.55-1.73	0.93
Allele 0 = C	HFS	879	0.076	0.14	1.32	0.88-1.97	0.18
CES2 intronic SNP rs11568311	GLOBAL	442	0.069	0.41	1.16	0.66-2.04	0.61
Allele 1 = A	DIARRHOEA	442	0.069	0.41	0.92	0.36-2.36	0.86
Allele 0 = G	HFS	441	0.069	0.41	1.38	0.75-2.53	0.30
CES2 intronic SNP rs11568314	GLOBAL	439	0.058	0.67	1.40	0.77-2.53	0.27
Allele 1 = T	DIARRHOEA	439	0.058	0.67	1.12	0.43-2.91	0.81
Allele 0 = A	HFS	438	0.058	0.67	1.60	0.84-3.03	0.15
CES2 intronic SNP rs2241409	GLOBAL	442	0.16	0.17	1.44	0.98-2.11	0.066
Allele 1 = T	DIARRHOEA	442	0.16	0.17	1.09	0.58-2.04	0.79
Allele 0 = C	HFS	441	0.16	0.17	1.58	1.03-2.41	0.035
DPYD *9A 85T>C	GLOBAL	838	0.23	0.31	0.91	0.71-1.17	0.47
Allele 1 = C	DIARRHOEA	838	0.23	0.31	0.87	0.59-1.26	0.45
Allele 0 = T	HFS	836	0.23	0.32	0.93	0.70-1.22	0.59
DPYD 496A>G	GLOBAL	923	0.09	0.49	0.89	0.63-1.25	0.50
Allele 1 = G	DIARRHOEA	923	0.09	0.49	0.95	0.56-1.60	0.84
Allele 0 = A	HFS	923	0.09	0.49	0.87	0.59-1.29	0.49
DPYD 1236G>A	GLOBAL	909	0.02	0.47	1.54	0.83-2.84	0.17
Allele 1 = A	DIARRHOEA	909	0.02	0.47	1.71	0.75-3.90	0.20
Allele 0 = G	HFS	909	0.02	0.47	1.40	0.71-2.76	0.33
DPYD *4A 1601G>A	GLOBAL	888	0.023	0.48	1.62	0.87-3.02	0.13
Allele 1 = A	DIARRHOEA	888	0.023	0.48	1.75	0.76-4.00	0.19
Allele 0 = G	HFS	886	0.023	0.48	0.97	0.46-2.05	0.94
DPYD *5 1627A>G	GLOBAL	907	0.20	0.44	1.02	0.80-1.31	0.87
Allele 1 = G	DIARRHOEA	907	0.20	0.44	1.06	0.73-1.53	0.77
Allele 0 = A	HFS	907	0.20	0.44	1.14	0.86-1.50	0.36
DPYD *2A IVS14+1G>A	GLOBAL	905	0.004	0.91	2.78	0.62-12.5	0.18
Allele 1 = A	DIARRHOEA	905	0.004	0.91	1.41	0.17-11.8	0.75
Allele 0 = G	HFS	903	0.004	0.91	2.67	0.59-12.0	0.20
DPYD *6 2194G>A	GLOBAL	912	0.032	0.27	1.06	0.62-1.84	0.82
Allele 1 = A	DIARRHOEA	912	0.032	0.27	1.78	0.88-3.56	0.10
Allele 0 = G	HFS	912	0.032	0.27	0.99	0.53-1.84	0.97
DPYD 2846A>T	GLOBAL	881	0.006	0.85	9.35	2.01-43.4	0.0043
Allele 1 = T	DIARRHOEA	881	0.006	0.85	3.14	0.82-11.9	0.093
Allele 0 = A	HFS	879	0.006	0.85	1.31	0.35-4.96	0.69

Table 3.6: Part 1 of 2. QUASAR2 capecitabine-related grade 012v34 toxicity results, under an allelic model, from counts. Previously studied putative 5-FU toxicity variants were analysed.

Polymorphism	Toxicity	N	MAF	HWE	OR	95% CI	p-value
MTHFR 677C>T	GLOBAL	927	0.32	0.66	1.10	0.89-1.35	0.39
Allele 1 = T	DIARRHOEA	927	0.32	0.66	1.09	0.79-1.49	0.60
Allele 0 = C	HFS	925	0.32	0.65	1.07	0.85-1.35	0.58
MTHFR 1298A>C	GLOBAL	917	0.33	0.31	0.93	0.75-1.14	0.47
Allele 1 = C	DIARRHOEA	917	0.33	0.31	0.92	0.67-1.27	0.62
Allele 0 = A	HFS	917	0.33	0.31	1.00	0.79-1.26	0.99
TYMP intronic SNP rs470119	GLOBAL	927	0.39	0.72	0.85	0.70-1.04	0.12
Allele 1 = A	DIARRHOEA	927	0.39	0.72	1.38	1.02-1.85	0.037
Allele 0 = G	HFS	925	0.39	0.73	0.79	0.63-0.99	0.044
TYMP S471L	GLOBAL	857	0.14	0.92	0.79	0.58-1.07	0.12
Allele 1 = T	DIARRHOEA	857	0.14	0.92	1.15	0.74-1.78	0.53
Allele 0 =C	HFS	855	0.14	0.93	0.73	0.52-1.04	0.082
TYMS 5'VNTR 2R/3R	GLOBAL	918	0.47	0.02	1.48	1.22-1.80	0.000079
Allele 1 = 2R	DIARRHOEA	918	0.47	0.02	1.29	0.96-1.74	0.093
Allele 0 = 3R	HFS	916	0.47	0.02	1.44	1.15-1.79	0.0013
TYMS 3'UTR 6bp ins-del	GLOBAL	474	0.31	0.27	1.67	1.23-2.22	0.00084
Allele 1 = ins	DIARRHOEA	474	0.31	0.27	1.49	0.94-2.38	0.085
Allele 0 = del	HFS	473	0.31	0.26	1.47	1.06-2.08	0.021
UMPS (OPRT) 638G>C	GLOBAL	927	0.20	0.42	1.09	0.86-1.40	0.47
Allele 1 = C	DIARRHOEA	927	0.20	0.42	0.83	0.56-1.23	0.36
Allele 0 = T	HFS	925	0.20	0.43	1.23	0.94-1.60	0.14

though the precise genetic model remained undefined. In order to estimate the per-allele effect across this haplotype, as would be clinically useful for patients who could not be phased, we performed a logistic regression analysis on the number of harmful alleles held by each patient (0 to 4) and found this "*TYMS* score test" to perform well for both global 012v34 toxicity (OR=1.38 per allele, p=0.00031) and decently for the common individual phenotypes diarrhoea grade 012v34 (OR=1.24 per allele, p=0.096) and HFS grade 012v34 (OR=1.31 per allele, p=0.0063).

3.10 Capecitabine Monotherapy Meta-Analysis Results

We then analysed the QUASAR2 results in conjunction with the data collected from other capecitabine monotherapy patients. Four studies, each comprising one patient set, contributed to the analysis (31; 115; 162; 175), in total contributing 395 patients – a minority relative to the 927 patients contributed by QUASAR2. Fifteen variants were analysed in at least two patient sets (including QUASAR2) and data for these SNPs were meta-analysed under an allelic model for global 5-FU-related toxicity, diarrhoea and HFS, all for grade 012v34 toxicity, as shown in Table 3.7.

Only one variant, the *TYMS* 5'VNTR 2R allele, was significant at the p<0.0033 FDR-adjusted threshold for the capecitabine monotherapy meta-analyses (OR=1.36, p=0.00028). This effect was largely driven by the QUASAR2 analysis, however, and results were non-significant but in the same direction when excluding QUASAR2 (OR=1.09, p=0.27). Results provided little support for the 3'UTR polymorphism (including QUASAR2: OR=1.35, p=0.012; excluding QUASAR2: OR=0.94, p=0.74), though this analysis showed evidence of inter-study heterogeneity resulting from the Largillier et al study (n=87). Raw data were available from the two largest studies (QUASAR2 and Caronia et al) for the 5'VNTR and 3'UTR variants, and the haplotype/independence analyses performed in the QUASAR2 dataset were repeated (n=602). In a logistic model for global grade 012v34 adjusted for study, again neither variant remained significant when including the other and again the 3'UTR (OR=1.42, p=0.017) was more significant than the 5'VNTR (OR=1.27, p=0.068). A similar pattern of increased effect size was observed across the haplotypes, in order of expected harmful effect based on single variant results (OR: 3R/del=1.00, 2R/del=1.25, 3R/ins=1.41, 2R/ins=1.79; p=0.0020), though the effect was less pronounced than for QUASAR2 alone. As with QUASAR2 alone, the incremental effect

Polymorphism	Toxicity	OR	95% CI	p-value	S	N	p-het
CDA -451C>T	GLOBAL	1.00	0.83-1.22	0.96	2	1035	0.0077
Allele 1 = T	DIARRHOEA	1.05	0.79-1.41	0.73	2	1035	0.54
Allele 0 = C	HFS	0.91	0.73-1.14	0.42	2	1035	0.006
CDA*2 79A>C	GLOBAL	1.03	0.86-1.24	0.71	3	1177	0.58
Allele 1 = C	DIARRHOEA	1.16	0.88-1.53	0.30	3	1177	0.11
Allele 0 = A	HFS	0.91	0.74-1.12	0.38	3	1175	0.42
CDA 943insC	GLOBAL	0.81	0.56-1.16	0.24	2	252	0.61
Allele 1 = insC	DIARRHOEA	0.72	0.38-1.37	0.32	2	252	0.41
Allele 0 = delC	HFS	0.81	0.53-1.23	0.33	2	252	0.064
CES2 823 (830) C/G 5'UTR	GLOBAL	1.19	0.88-1.61	0.25	3	1134	0.96
Allele 1 = G	DIARRHOEA	0.91	0.55-1.49	0.70	3	1134	0.72
Allele 0 = C	HFS	1.30	0.94-1.81	0.12	3	1132	0.97
CES2 intronic SNP rs11568314	GLOBAL	1.17	0.69-1.97	0.56	2	569	0.25
Allele 1 = T	DIARRHOEA	0.91	0.39-2.16	0.83	2	569	0.44
Allele 0 = A	HFS	1.33	0.76-2.32	0.32	2	568	0.29
CES2 intronic SNP rs11568311	GLOBAL	1.14	0.70-1.86	0.59	2	572	0.92
Allele 1 = A	DIARRHOEA	1.04	0.49-2.23	0.92	2	572	0.63
Allele 0 = G	HFS	1.20	0.71-2.04	0.50	2	571	0.41
CES2 intronic SNP rs2241409	GLOBAL	1.15	0.83-1.59	0.40	2	572	0.046
Allele 1 = T	DIARRHOEA	0.88	0.52-1.49	0.62	2	572	0.25
Allele 0 = C	HFS	1.29	0.90-1.83	0.16	2	571	0.11
DPYD *2A IVS14+1G>A	GLOBAL	3.02	0.78-11.7	0.11	2	1035	0.83
Allele 1 = A	DIARRHOEA	3.14	0.71-13.9	0.13	2	1035	0.18
Allele 0 = G	HFS	1.98	0.52-7.54	0.32	2	1033	0.46
MTHFR 677C>T	GLOBAL	1.09	0.89-1.32	0.40	3	1061	0.74
Allele 1 = T	DIARRHOEA	1.03	0.76-1.39	0.85	3	1068	0.45
Allele 0 = C	HFS	1.05	0.85-1.32	0.64	3	1066	0.95
MTHFR 1298A>C	GLOBAL	0.97	0.79-1.18	0.74	3	1051	0.20
Allele 1 = C	DIARRHOEA	0.99	0.73-1.34	0.97	3	1059	0.27
Allele 0 = A	HFS	1.04	0.84-1.30	0.71	3	1058	0.35
TYMP S471L	GLOBAL	0.79	0.59-1.05	0.11	2	985	0.99
Allele 1 = T	DIARRHOEA	1.08	0.71-1.65	0.72	2	985	0.34
Allele 0 = C	HFS	0.72	0.52-1.01	0.058	2	987	0.80
TYMP intronic rs470119	GLOBAL	0.85	0.70-1.02	0.083	2	1056	0.88
Allele 1 = A	DIARRHOEA	1.22	0.92-1.61	0.17	2	1056	0.033
Allele 0 = G	HFS	0.80	0.65-0.99	0.038	2	1054	0.80
TYMS 5'VNTR 2R/3R repeat	GLOBAL	1.36	1.15-1.60	0.00028	5	1300	0.17
Allele 1 = 2R	DIARRHOEA	1.12	0.87-1.45	0.38	5	1309	0.29
Allele 0 = 3R	HFS	1.33	1.10-1.60	0.0029	5	1306	0.23
TYMS 5'VNTR 3R USF G/C SNP	GLOBAL	1.12	0.65-1.95	0.68	3	260	0.23
Allele 1 = 3/4 USF sites	DIARRHOEA	1.08	0.50-2.34	0.84	3	267	0.88
Allele 0 = 0/1/2 USF sites	HFS	1.27	0.69-2.34	0.44	3	267	0.45
TYMS 3'UTR 6bp ins-del	GLOBAL	1.35	1.07-1.71	0.012	4	738	0.025
Allele 1 = ins	DIARRHOEA	0.99	0.70-1.39	0.94	4	745	0.006
Allele 0 = del	HFS	1.43	1.09-1.87	0.0091	4	743	0.34

Table 3.7: Meta-analysed capecitabine monotherapy-related grade 012v34 toxicity results, under an allelic model, from counts. Fifteen variants were selected for analysis based on previous study in at least two of five collected capecitabine monotherapy studies. "S"=studies; "N"=patients.

was not observed for diarrhoea and HFS. When we applied the same *TYMS* score test to each of these two studies and meta-analysed the results, we observed improved significance relative to QUSAR2 for global (OR=1.33 per allele, p=0.00018) and HFS (OR=1.29 per allele, p=0.0030).

Unfortunately, no other capecitabine monotherapy data were available for the *DPYD* 2846 variant that showed strong significance in the QUASAR2 analysis, though the effect size for the *2A variant increased somewhat.

3.11 Infusional and Bolus 5-FU Monotherapy Meta-Analysis Results

With no convincing precedent in the literature showing that data from patients can be jointly analysed across 5-FU regimens, we analysed the additional 5-FU monotherapy regimens (infusional and bolus) separately from each other and from capecitabine. Later, we considered whether results and data from these regimens could be further combined, as discussed in Section 3.15.

Data from three bolus and four infusional 5-FU patient sets were available. Pooling data across the patient sets for each regimen, we determined that the toxicity profile of these regimens differed from capecitabine, as previously shown in Table 3.4. For infusional 5-FU, the only common (ie, analysed) phenotype was diarrhoea, whereas for bolus 5-FU, common phenotypes included diarrhoea, mucositis/stomatitis and neutropaenia. Eight variants had data from at least two patient sets from the same regimen, and these were meta-analysed, as shown in Table 3.8. As a point of comparison, where data for a polymorphism were available from a single patient set for one regimen but meta-analysed in the other regimen, the summary statistics are provided for both infusional and bolus 5-FU. Additional variants were analysed for a single report (results not shown).

		Infusional 5-FU monotheray					Bolus 5-FU monotherapy				
Polymorphism	**Toxicity**	**OR**	**95% CI**	**p-value**	**S**	**N**	**OR**	**95% CI**	**p-value**	**S**	**N**
DPYD 496A>G	GLOBAL	0.53	0.09-3.10	0.48	1	36	1.25	0.78-2.03	0.35	2	379
Allele 1 = G	DIARRHOEA	0.37	0.04-3.39	0.38	1	36	1.40	0.81-2.42	0.22	2	379
Allele 0 = A	MUCOSITIS/STOMATITIS						0.92	0.47-1.79	0.80	2	379
	NEUTROPAENIA						1.17	0.52-2.60	0.70	2	379
DPYD *5 1627A>G	GLOBAL	0.68	0.27-1.75	0.43	1	117	0.71	0.49-1.02	0.062	2	452
Allele 1 = G	DIARRHOEA	0.67	0.22-2.06	0.48	1	117	0.79	0.49-1.28	0.34	2	452
Allele 0 = A	MUCOSITIS/STOMATITIS						0.49	0.27-0.88	0.018	2	452
	NEUTROPAENIA						0.73	0.35-1.51	0.39	2	452
DPYD *2A IVS14+1G>A	GLOBAL	6.71	1.66-27.1	0.0075	2	732	3.84	0.95-15.6	0.059	1	338
Allele 1 = A	DIARRHOEA	7.71	1.61-36.9	0.011	2	720	1.47	0.18-12.2	0.72	1	338
Allele 0 = G	MUCOSITIS/STOMATITIS						7.15	1.75-29.1	0.0061	1	338
	NEUTROPAENIA						12.90	3.13-53.3	0.0004	1	338
MTHFR 677C>T	GLOBAL	0.92	0.70-1.22	0.57	3	1129	0.79	0.62-1.00	0.05	3	788
Allele 1 = T	DIARRHOEA	0.88	0.58-1.33	0.54	3	1112	0.75	0.54-1.06	0.10	3	788
Allele 0 = C	MUCOSITIS/STOMATITIS						0.99	0.73-1.33	0.93	3	788
	NEUTROPAENIA						0.60	0.37-0.97	0.039	3	788
MTHFR 1298A>C	GLOBAL	1.01	0.67-1.53	0.94	2	474	1.03	0.81-1.29	0.83	3	786
Allele 1 = C	DIARRHOEA	0.88	0.52-1.49	0.64	2	474	1.01	0.74-1.39	0.94	3	786
Allele 0 = A	MUCOSITIS/STOMATITIS						0.98	0.73-1.32	0.88	3	786
	NEUTROPAENIA						1.36	0.90-2.06	0.14	3	786
TYMS 5'VNTR 2R/3R repeat	GLOBAL	1.45	1.13-1.85	0.0035	4	1185	1.40	1.04-1.9	0.028	2	466
Allele 1 = 2R	DIARRHOEA	1.45	1.01-2.08	0.042	4	1168	1.65	1.05-2.58	0.03	2	466
Allele 0 = 3R	MUCOSITIS/STOMATITIS						1.31	0.93-1.86	0.13	2	466
	NEUTROPAENIA						1.31	0.79-2.19	0.29	2	466
TYMS 5'VNTR 3R USF G/C SNP	GLOBAL	1.07	0.45-2.58	0.87	2	204	0.90	0.59-1.37	0.62	2	427
Allele 1 = 3/4 USF sites	DIARRHOEA	1.42	0.48-4.23	0.52	2	204	1.10	0.63-1.95	0.73	2	427
Allele 0 = 0/1/2 USF sites	MUCOSITIS/STOMATITIS						0.89	0.52-1.53	0.68	2	427
	NEUTROPAENIA						0.70	0.28-1.74	0.44	2	427
TYMS 3'UTR 6bp ins-del	GLOBAL	1.16	0.88-1.53	0.31	4	1125	1.98	1.15-3.40	0.00038	2	456
Allele 1 = ins	DIARRHOEA	1.27	0.84-1.94	0.26	4	1109	1.88	1.33-2.66	0.014	2	456
Allele 0 = del	MUCOSITIS/STOMATITIS						2.03	1.34-3.08	0.00086	2	456
	NEUTROPAENIA						1.28	0.73-2.25	0.39	2	456

Table 3.8: Meta-analysed infusional and bolus monotherapy-related grade 012v34 toxicity results, under an allelic model, from counts. Eight variants were selected for analysis based on previous study in at least two of three collected bolus monotherapy or two of four collected infusional monotherapy patient sets. "S"=studies; "N"=patients.

Only one variant met the p<0.0048 significance threshold in each regimen for global grade 012v34 toxicity: the *TYMS* 5'VNTR 2R for infusional monotherapy (OR=1.45, p=0.0035) and the *TYMS* 3'UTR ins for bolus monotherapy (OR=1.98, p=0.00038). The infusional and bolus results complicated the understanding of the interaction between these *TYMS* variants identified in capecitabine. Whereas for capecitabine the 3'UTR polymorphism remained significant when adjusted for the 5'VNTR in the logistic model, the opposite was true for infusional 5-FU (n=957 from three studies), with the 5'VNTR showing significance when adjusted for the 3'UTR (OR=1.53, p=0.0040) and the 3'UTR remaining non-significant (OR=0.94, p=0.68). Further, the 2R/del haplotype (observed as second least harmful for capecitabine) was observed as the most harmful haplotype for infusion 5-FU patients. For bolus, analyses were closer, but not identical, to the capecitabine results, with the 3'UTR showing more significance when adjusted for the 5'VNTR and the 3R/ins haplotype acting as the most harmful (as opposed to second most harmful for capecitabine). Applying the *TYMS* score test separately to three infusional (n=957) and one bolus (n=325) 5-FU patient sets, results were only nominally significant (OR=1.21 per allele, p=0.035; OR=1.35 per allele, p=0.014). A table summarising the *TYMS* haplotype and other top results from monotherapy and combination regimens is shown in Section 3.14.

The *DPYD* *2A variant approached significance in the infusional 5-FU analysis for global grade 012v34 toxicity (OR=6.71, p=0.0075); the result is from a single small study, suggesting that further analysis in this regimen might lead to formal significance. Further, this variant was significant for neutropaenia – the most clinically important phenotype – in the bolus regimen (OR=12.90, p=0.00040). Of note, the *DPYD* 2846 variant did not have sufficient data for meta-analysis in either regimen, but in a single-patient-set analysis in each regimen from Schwab et al (173), the A allele was of large effect size and in the same direction as for QUASAR2 (infusional:

OR=3.09, p=0.36; bolus: OR=19.3, p=0.056).

3.12 Combination Therapy 5-FU Meta-Analysis Results

Five variants were available for analysis in 5-FU-based combination regimens (FOLFOX, XELOX, FOLFIRI, IFL, FLIRI), of which FOLFOX (infusional 5-FU plus oxaliplatin) had data from as many as five studies (n>800) for some polymorphisms/phenotypes (Table 3.9). No variants reached either the p<0.0048 significance or p<0.05 suggestive significance threshold, including the *TYMS* 5'VNTR and 3'UTR polymorphisms, whose results from individual patient sets are shown in forest plots in Figure 3.3. These variants did not trend in the same direction of effect that was observed in all of the monotherapy regimens, as a whole suggesting that polymorphisms that predict toxicity in 5-FU monotherapy regimens are not necessarily generalisable to the 5-FU combination therapy regimens. Of note, no data were available for meta-analysis for either the *DPYD* 2846 or *2A variant, meaning that there is no direct evidence against functional *DPYD* variants having an impact on toxicity from 5-FU combination regimens.

		Infusional + oxaliplatin (FOLFOX)				Infusional + irinotecan (FOLFIRI)				Bolus + irinotecan			
Polymorphism	**Toxicity**	**OR**	**95% CI**	**p-value**	**S**	**N**	**OR**	**95% CI**	**p-value**	**S**	**N**	**OR**	**95% CI**	**p-value**	**S**	**N**
MTHFR 677C>T	GLOBAL	1.13	0.87-1.46	0.38	4	688	1.18	0.69-2.00	0.55	2	226	1.34	0.74-2.45	0.34	2	178
Allele 1 = T	DIARRHOEA	1.31	0.90-1.91	0.15	5	806										
Allele 0 = C	NEUTROPAENIA	1.03	0.79-1.34	0.83	4	692	1.38	0.72-2.65	0.33	2	226	0.93	0.53-1.62	0.8	2	178
	THROMBOCYTOPAENIA	0.87	0.52-1.46	0.60	4	693										
MTHFR 1298A>C	GLOBAL	0.96	0.71-1.30	0.80	3	528	2.53	0.82-7.81	0.11	1	65	0.53	0.27-1.05	0.069	2	171
Allele 1 = C	DIARRHOEA	0.96	0.59-1.57	0.89	3	535										
Allele 0 = A	NEUTROPAENIA	1.07	0.80-1.43	0.63	3	534	5.40	0.55-53.45	0.15	1	65	0.92	0.51-1.63	0.77	2	171
	THROMBOCYTOPAENIA	1.13	0.68-1.89	0.63	3	534										
TYMS 5'VNTR 28bp repeat	GLOBAL	0.87	0.68-1.11	0.26	5	726	0.83	0.59-1.18	0.31	2	309					
Allele 1 = 2R	DIARRHOEA	0.94	0.65-1.37	0.75	5	728						1.49	0.78-2.84	0.22	1	103
Allele 0 = 3R	NEUTROPAENIA	0.91	0.71-1.16	0.44	5	730	1.32	0.87-2.02	0.20	2	309	0.70	0.40-1.22	0.21	1	103
	THROMBOCYTOPAENIA	1.22	0.77-1.93	0.40	5	731										
TYMS 5'VNTR 3R USF G/C SNP	GLOBAL	0.89	0.53-1.51	0.67	3	309										
Allele 1 = 3/4 USF sites	DIARRHOEA	1.75	0.72-4.24	0.21	3	316										
Allele 0 = 0/1/2 USF sites	NEUTROPAENIA	0.79	0.45-1.40	0.42	3	315										
	THROMBOCYTOPAENIA	0.78	0.26-2.33	0.66	3	315										
TYMS 3'UTR 1494indel6bp	GLOBAL	1.03	0.79-1.34	0.80	4	686	1.10	0.76-1.59	0.62	2	305					
Allele 1 = ins	DIARRHOEA	0.82	0.54-1.24	0.34	4	688						0.59	0.30-1.15	0.12	1	103
Allele 0 = del	NEUTROPAENIA	1.09	0.84-1.42	0.51	4	690	1.24	0.78-1.96	0.36	2	305	0.81	0.45-1.47	0.49	1	103
	THROMBOCYTOPAENIA	1.36	0.79-2.34	0.27	4	691										

Table 3.9: Meta-analysed 5-FU-based combination therapy-related grade 012v34 toxicity results, under an allelic model, from counts. Five variants were selected for analysis based on previous study in at least two patient sets of the same regimen and single set results are reported in other regimens where available for the same polymorphisms. One study (n=47) was available for capecitabine + oxaliplatin (XELOX); data not shown but all results were at p>0.15 and no "global" measure was available. "S"=studies; "N"=patients.

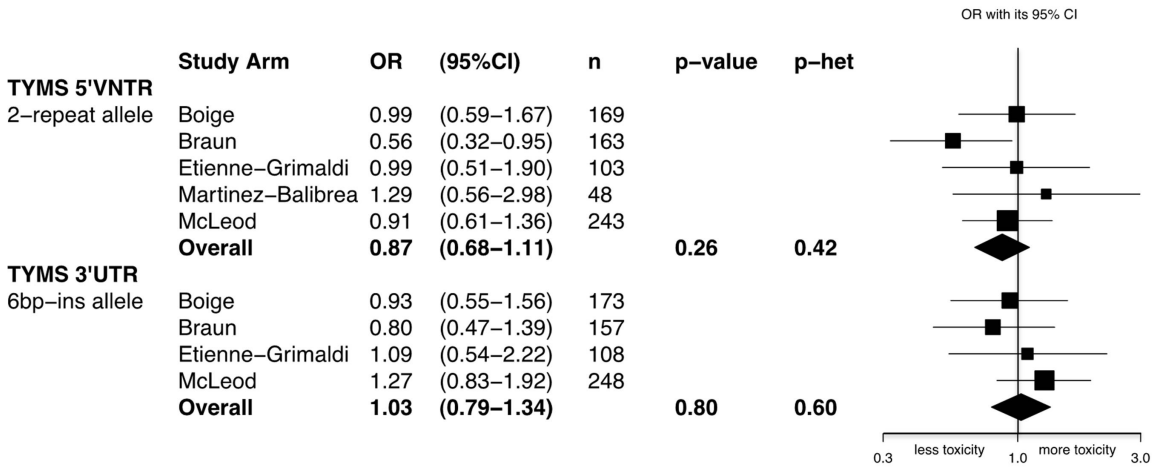


Figure 3.3: Forest plot showing results from the individual patient sets included in the meta-analysis of the *TYMS* 5'VNTR and 3'UTR polymorphisms, for patients on the FOLFOX regimen. No significant difference from "no effect" was observed. All analyses are for the binary grade 012v34 outcome.

3.13 DPYD Deficiency Syndrome Literature Review and *DPYD* Combined Analysis

The *DPYD* gene has a very strong candidate basis, given its role in catabolising approximately 80% of injected 5-FU prior to activation, lowering the effective dose for most patients (84). Among candidate genes, *DPYD* has been studied the most and positive associations have been made with many studied coding variants. Most of these associations have been weak (as shown in the previous meta-analysis sections), however, with the exception of some rare coding or splice variants such as 2846 and 2A. With a QUASAR2 minor allele frequency of 0.006 and 0.004, respectively, the variants require at least 1,000 patients in order to show significance. Given that homozygote and compound heterozygote carriers of these alleles will develop DPYD Deficiency Syndrome (OMIM 274270) (203; 206), negative selection will exist for these variants and thus we would expect that additional similar variants that might exist would be at an equally rare frequency in the population, if not rarer, given their absence in QUASAR2.

Inclusion	Polymorphism	rsID	Patient Ethnicity	Affected Alleles	MAF	SIFT (score)	Polyphen	Structural Impact	Homology	Citation
Lead	*DPYD* 2846T>A; D949V	rs67376798	Dutch	CH with *DPYD**2A	0.0055	damaging (0)	0.98	Cofactor binding or electron transport	E, D, M, F, N	2002
No	*DPYD**9A; 85T>C; C29R	rs1801265	Italian	Homozygous	0.29	tolerated (0.38)	0.66	Predicted to have no impact	None	1999
Yes	*DPYD* 257C>T; Pro86Leu	not listed	Eritrean-Ethiopian	Homozygous	no data	no data	no data	Cofactor binding or electron transport	E, D, M, F, N	2002
Yes	*DPYD**7; 295–298delTCAT	rs72549309	3 CEU patients	Homozygous	no data	no data	no data	Frameshift		1999
Yes	*DPYD* 601A>C; Ser201Arg	rs72549308	Scottish	CH with *DPYD**2A	no data	damaging (0)	1.00	Cofactor binding or electron transport	E, D, M, F, N	2002
Yes	*DPYD* 632A>G; Tyr211Cys	rs72549307	Turkish	Homozygous	no data	damaging (0)	1.00	Destabilizing effects	E, D, M, F, N	2002
Yes	*DPYD* 731A>C; Glu244Val	not listed	not reported	CH with *DPYD**2A	no data	tolerated (0.09)	0.94	Cofactor binding or electron transport	E, D, M	2005
Yes	*DPYD**8; 703C>T; R235W	rs1801266	Turkish	CH with 1897delC	0.0001	damaging (0)	no data	Not modelled	E, D, M, F, N	1999
Yes	*DPYD* 1039-1042delTG	not listed	Greek-Ukrainian	CH with I560S	no data	no data	no data	Frameshift		2002
No	*DPYD* 1108A>G; Ile370Val	rs72549305	Asian	Homozygous	no data	tolerated (0.07)	0.02	Predicted to have no impact	E, D, M, F, N	2002
Yes	*DPYD* IVS11+1G>T	not listed	not reported	CH with *DPYD**2A	no data	no data	no data	Deletion deletion of residues 400–446		2005
Yes	*DPYD* 1475C>T; Ser492Leu	rs72549304	Pakistani	Homozygous	no data	damaging (0)	0.74	Cofactor binding or electron transport	E, D, M, F, N, Y	2002
Yes	*DPYD* 1651G>A; Ala551Thr	not listed	not reported	Homozygous	no data	damaging (0)	1.00	Cofactor binding or electron transport	E, D, M, F, N, Y	2005
Yes	*DPYD* *13; 1679T>G; I560S	rs55886062	Greek-Ukrainian	CH with 1039-1042delTG	0.0006	damaging (0)	0.98	Destabilizing effects	E, D, M, F	2002
Yes	*DPYD**2A; IVS14+1G>A Del	rs3918290	9 CEU patients	Homozygous	0.004	no data	no data	Exon skipping		1999
Yes	*DPYD**3; 1897delC	rs72549303	Turkish	Homozygous	no data	no data	no data	Frameshift		1999
Yes	*DPYD**9B; 2657G>A; R886H	rs1801267	Turkish	Homozygous	no data	damaging (0)	0.98	Not modelled	E, D, M, F, N	1999
Yes	*DPYD* 2933A>G; His978Arg	rs72547601	Pakistani	Homozygous	no data	damaging (0)	1.00	Cofactor binding or electron transport	E, D, M, F, N	2002
Yes	*DPYD**10; 2983G>T; V995F	rs1801268	Dutch	Homozygous	no data	damaging (0)	0.96	Not modelled	E, D, M, F, N	1999

Table 3.10: Putative DPYD Deficiency Syndrome causal variants are listed and annotated. Of the 19 variants previously suggested to result in the syndrome when held as a homozygote or compound heterozygote with an additional causal variant, 17 are likely causal. Given the significant association observed between the *DPYD* 2846 variant (as a lone heterozygote) and capecitabine-related toxicity, we assume that these 17 variants similarly affect drug toxicity, even though most are too rare to analyse by association study. Original citations all from Van Kuilenburg et al (203; 205; 206). CH=compound heterozygote. MAF according to largest CEU population with data from either dbSNP, Exome Variant Server, or QUASAR2 (if no other data available). Structural impact according to van Kuilenburg et al 2002 and 2005. Sequence homology with elephant (E), dog (D), mouse (M), fly (F), nematode (N), and yeast (Y) identified using ClustalW.

An alternative means of studying the impact of rare *DPYD* variants on capecitabine/5-FU-related toxicity is to perform a functional analysis. One such experiment is an *in vivo* analysis of whether or not the rare variant results in DPYD Deficiency Syndrome when held as a homozygote or compound heterozygote. As reviewed in Table 3.10, the causal variants of DPYD Deficiency Syndrome have been tracked for 15 years by Van Kuilenburg et al (203; 205; 206). Importantly, the 2846 variant is common enough that we have been able to identify its impact on capecitabine-related toxicity by association study, as shown in Section 3.9, and this variant has also been shown to cause DPYD Deficiency Syndrome, specifically in a patient who was compound heterozygote with the *DPYD* *2A rare allele (203). Further functional annotation based on databases and from the original report suggests that the variant is highly conserved and likely impacts cofactor binding or electron transport in the protein. There are 19 additional variants that have been described as causing the syndrome, however, one is too common to be causal (*9A; 85T>C) and a second is predicted as unlikely to have a functional impact (Ile370Val; 1108A>G). By extension from 2846, then, we would expect that 16 previously identified additional rare *DPYD* variants would increase risk of toxicity from capecitabine and possibly other 5-FU regimens, including the *2A variant. As further evidence, both the *2A and I560S (1679T>G) variant have been shown to greatly reduce DPYD activity *in vitro* using human embryonic kidney and human colon carcinoma cell lines transfected with *DPYD* containing various mutations, whereas the more common mutations reported caused no reduction to activity (149).

In order to determine the effect size expected from variants in this functional class, we performed a combined assessment of the causal DPYD Deficiency Syndrome variants with data available in the meta-analysis (*DPYD* *2A and 2846T>A). This "group test" was performed by the same methods used in the analysis of the variants individually (ie, a 2-by-2 table of allele counts for low and high toxicity patients),

except that rare alleles from either variant contributed to an individual's count of harmful alleles (no patient had more than one rare allele). Results of this *DPYD* combined model are shown in Table 3.11 using data from QUASAR2 and Schwab et al (173), separately for each of the three 5-FU-based monotherapy regimens. We see that for the measure of global toxicity, for each of bolus, infusional and capecitabine, the significance of the group model is greater than that of either variant alone and formal significance is achieved for capecitabine. While this is due to an increased amount of data, significance would decrease if there were strong evidence that one variant did not impact toxicity. For example, this is seen for infusional 5-FU for diarrhoea. Thus, for global toxicity, we would expect that any of the 17 previously identified DPYD Deficiency Syndrome polymorphisms would increase risk about 5.5-fold for capecitabine patients (OR=5.51, p=0.0013) and similarly for patients on other regimens. As both patients who experienced treatment-related deaths in QUASAR2 held one of these rare alleles, it is additionally likely that these rare alleles shift the spectrum of toxicities experienced. Of note, the impact described for the previously identified variants would extend to any rare functional *DPYD* variant.

Toxicity	Study	Regimen	2846 A only	*2A A only	2846 A or *2A A allele
Global	Schwab	Bolus	19.34 (0.92-405.26), p=0.056	3.84 (0.95-15.55), p=0.059	5.85 (1.63-21.01), p=0.0068
	Schwab	Infusional	3.09 (0.28-34.43), p=0.36	6.36 (0.88-45.69), p=0.066	4.80 (1.06-21.82), p=0.042
	QUASAR2	Capecitabine	9.35 (2.01-43.41), p=0.0043	2.78 (0.62-12.48), p=0.18	5.51 (1.95-15.51), p=0.0013
Diarrhoea	Schwab	Bolus	2.17 (0.1-45.71), p=0.62	1.47 (0.18-12.19), p=0.72	1.10 (0.14-9.18), p=0.90
	Schwab	Infusional	1.42 (0.07-27.83), p=0.82	10.82 (1.5-78.3), p=0.018	4.30 (0.82-22.71), p=0.085
	QUASAR2	Capecitabine	3.14 (0.82-11.92), p=0.093	1.41 (0.17-11.75), p=0.75	2.48 (0.81-7.6), p=0.11
Muc/Stom	Schwab	Bolus	36.08 (1.72-758.5), p=0.021	7.15 (1.75-29.13), p=0.0061	11 (3.04-39.78), p=0.00030
Neutropenia	Schwab	Bolus	2.26 (0.11-47.63), p=0.60	12.92 (3.13-53.26), p=0.00040	8.6 (2.34-31.46), p=0.0012
HFS	QUASAR2	Capecitabine	1.31 (0.35-4.96), p= 0.70	2.67 (0.59-11.96), p=0.20	1.76 (0.66-4.71), p=0.26

Table 3.11: The effect of having either the *DPYD* *2A or 2846 rare allele was tested for each of the three 5-FU-based monotherapy regimens and compared to the effect observed from either of the variants alone. Increased significance suggests that, as a functional class, variants causal of DPYD Deficiency Syndrome increase risk of dose limiting levels of global 5-FU-related toxicity 5- to 6-fold.

3.14 Summary of Significant Findings

Of the 36 previously-assessed polymorphisms analysed, only four - *TYMS* 5'VNTR 2R/3R, *TYMS* 3'UTR 6bp ins-del, *DPYD* 2846T>A, and *DPYD* *2A - showed formal associations with global dose limiting toxicity. While formal significance ($p<0.0071$ for the QUASAR2 analysis, $p<0.0033$ for the capecitabine meta-analyses, and $p<0.0048$ for the non-capecitabine meta-analyses) was not reached in all regimens, these polymorphisms were all significant for either the capecitabine meta-analysis or the QUASAR2 analysis (our most powered analyses), and the meta-analysed result of the other 5-FU monotherapy regimens was in the same direction. As further discussed in the next section, this apparent consistency among monotherapy regimens led us to investigate whether information could be shared across regimens via formally combined analyses. In addition, rare functional *DPYD* variants, such as the 17 identified by literature review, appear to increase one's risk of toxicity.

Strikingly, no polymorphism was significant in any of the combination regimens, including for FOLFOX, which had over 600 patients for analysis. While some analyses suffered from low power that prevented us from excluding variants from having an effect on toxicity, other analyses were well powered. Power to detect an association for the common (MAF>8%) *DPYD* 1627A>G, *DPYD* 85T>C, *DPYD* 496A>G, *TYMS* 5'VNTR G>C, *MTHFR* 677C>T, *MTHFR* 1298A>C, *CDA* -451C>T, *CES2* 823C>G, and *TYMP* polymorphisms was approximately 75% to 100%, while rarer variants including *DPYD* 1601G>A, *DPYD* 1236G>A, *DPYD* 2194G>A, *CDA* 943insC, and most *CES2* polymorphisms could not be ruled out with confidence (approximately 20-40% power). Forest plots of monotherapy regimens for the four significant polymorphisms and *MTHFR* 1298 (a common variant previously named for an association with toxicity, but that was flat across our analysis) are shown in Figures 3.4, 3.5, 3.6, 3.7, and 3.8. Among other patterns, these figures show that more variants are significant for bolus 5-FU than for other regimens, possibly suggesting that either

the underlying mechanisms of toxicity differ among regimens or that bolus 5-FU acts as a harbinger of pharmacogenetic effect for other 5-FU regimens owing to its higher event rate.

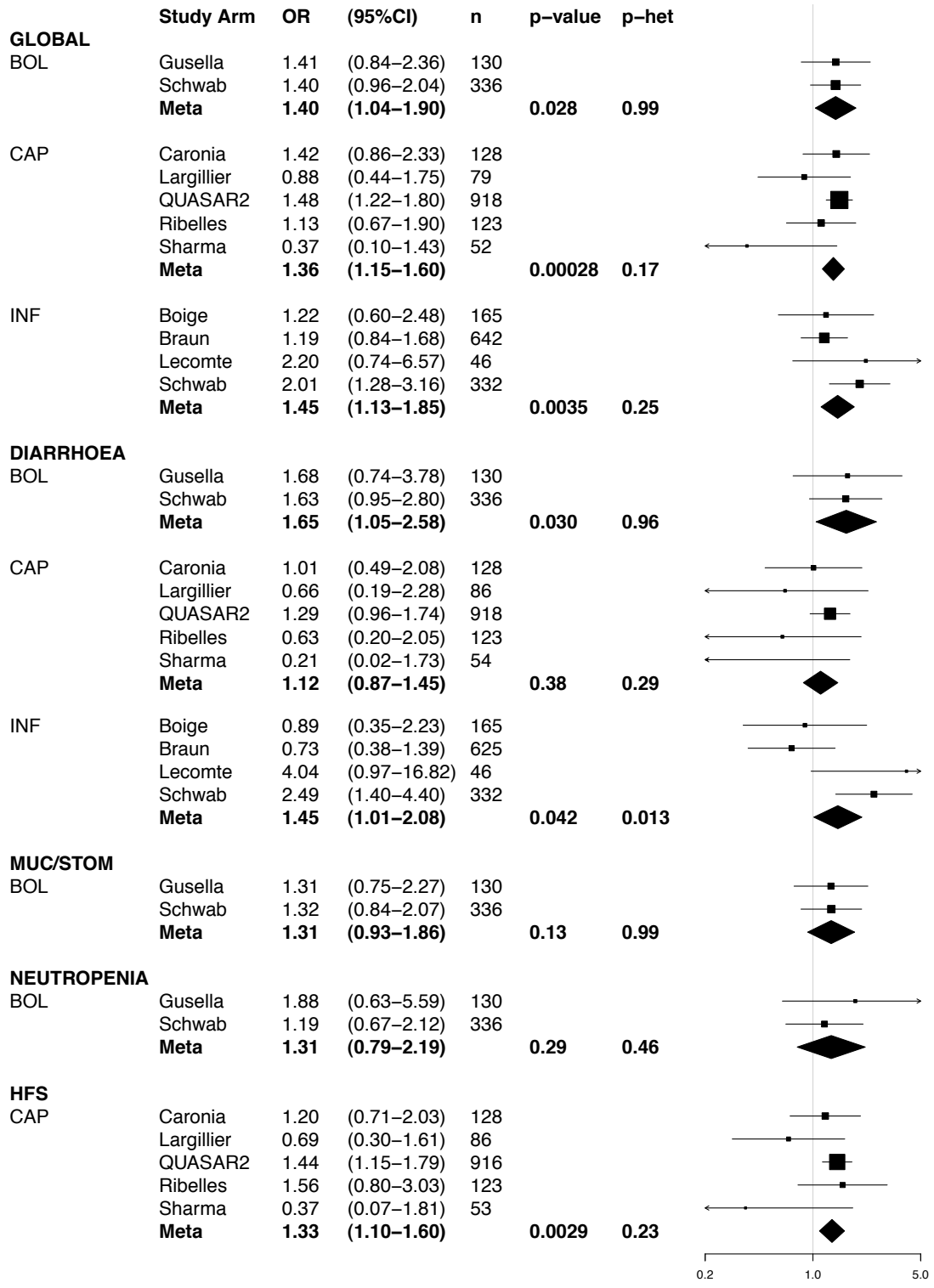


Figure 3.4: Forest plot showing results from the individual patient sets included in the meta-analyses of the *TYMS* 5'VNTR polymorphism 2R allele for patients on 5-FU monotherapy regimens. Results were significant for global toxicity among capecitabine and infusional 5-FU patients. The impact on global toxicity appears driven by both diarrhoea and rare phenotypes for bolus 5-FU and mostly driven by HFS for capecitabine. All analyses are for the binary grade 012v34 outcome.

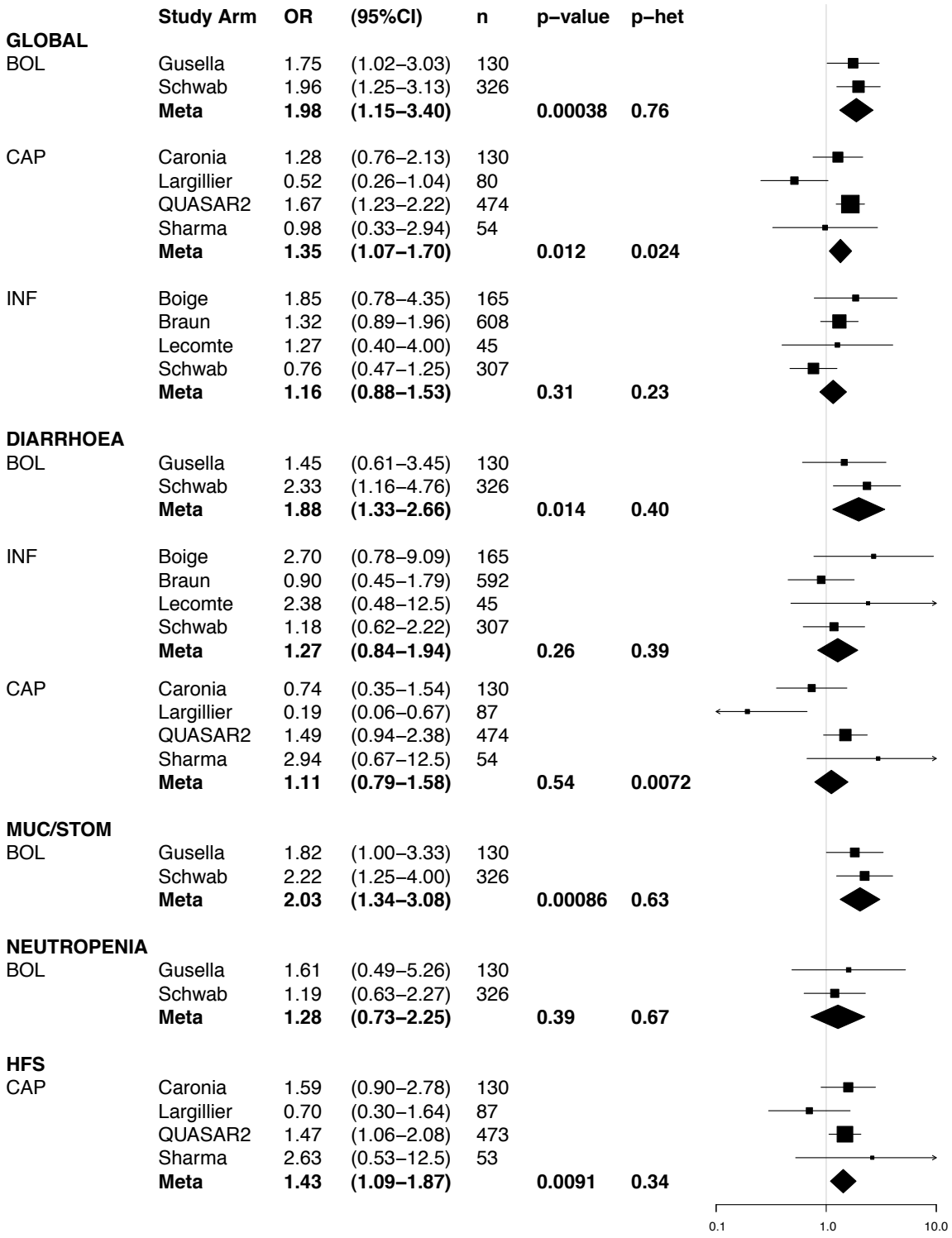


Figure 3.5: Forest plot showing results from the individual patient sets included in the meta-analyses of the *TYMS* 3'UTR polymorphism insertion allele for patients on 5-FU monotherapy regimens. Results were significant for global toxicity among bolus 5-FU patients, for which the impact appears driven by both diarrhoea and rare phenotypes. Results for the *TYMS* score test combining both the 3'UTR and 5'VNTR polymorphism, however, were significant for capecitabine patients, as discussed in this section. All analyses are for the binary grade 012v34 outcome.

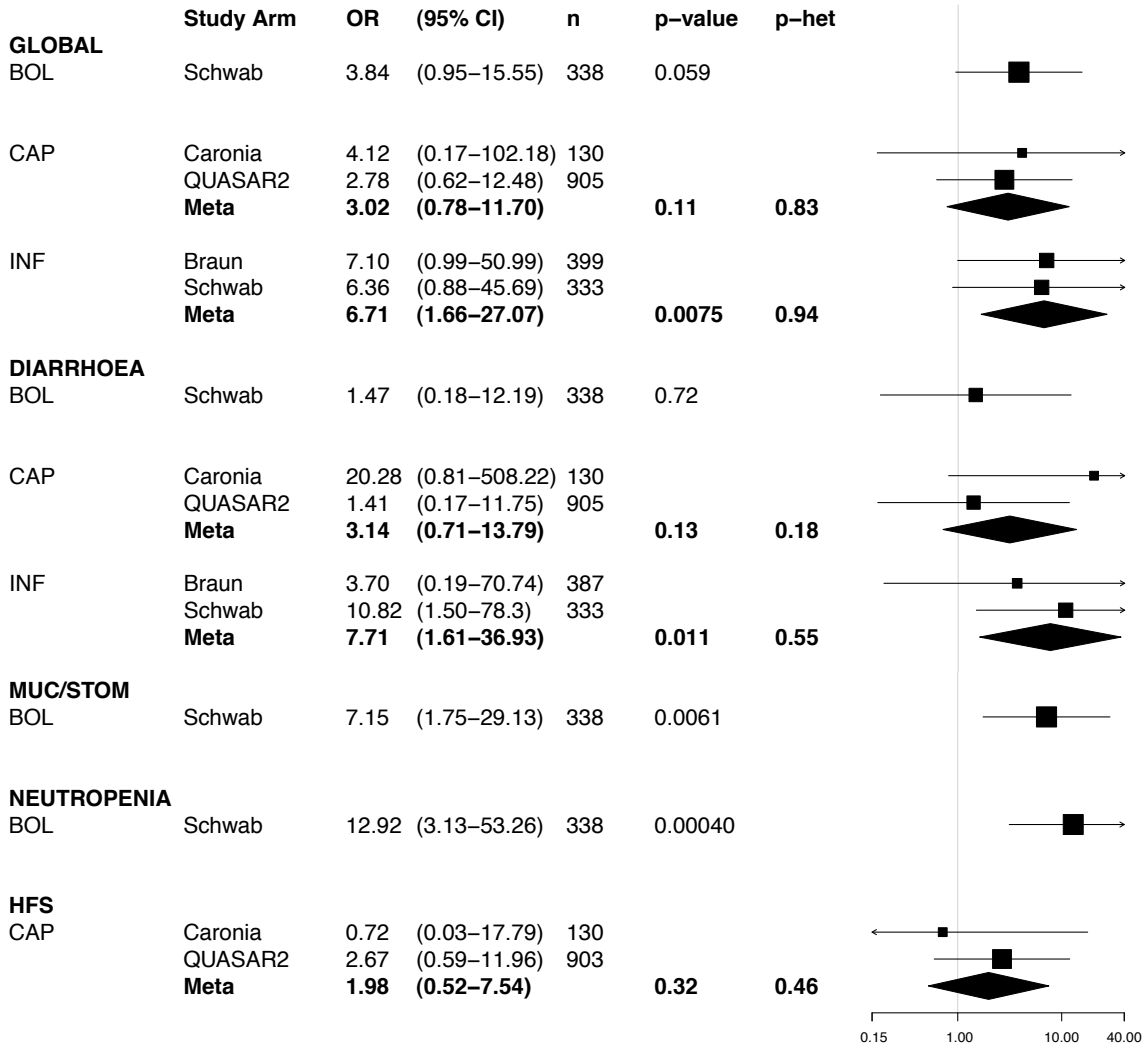


Figure 3.6: Forest plot showing results from the individual patient sets included in the meta-analyses of the *DPYD* *2A polymorphism A allele for patients on 5-FU monotherapy regimens. Results for this rare variant were not formally significant for global toxicity for any regimen, though the direction of effect was consistent and the *DPYD* group test with the 2846 variant was significant among capecitabine patients and reached suggestive significance among bolus and infusional 5-FU patients, as described in this section. All analyses are for the binary grade 012v34 outcome.

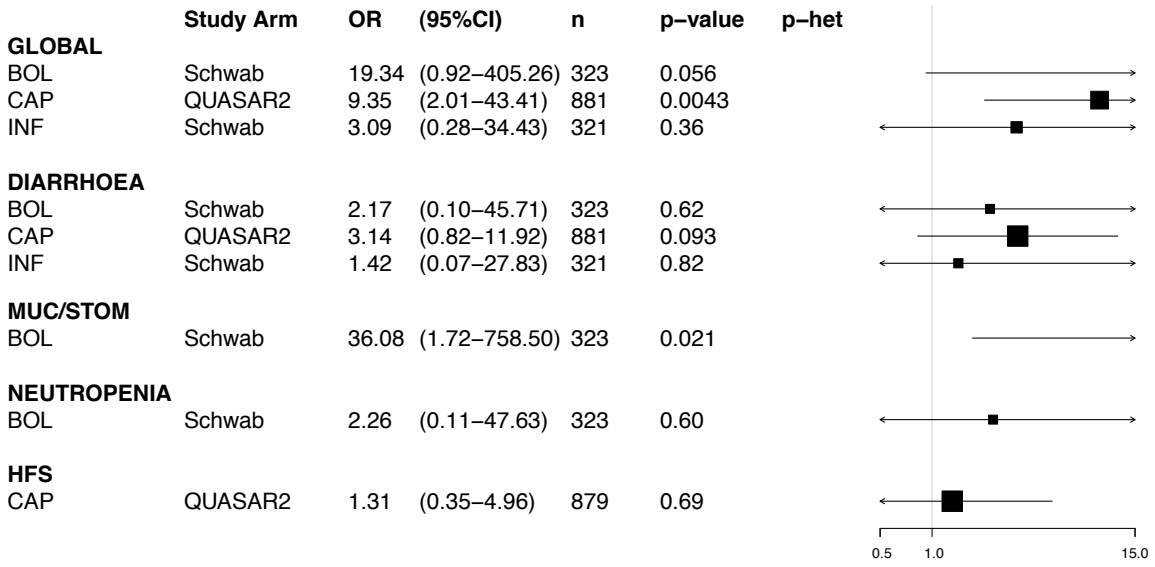


Figure 3.7: Forest plot showing results from the individual patient sets included in the meta-analyses of the *DPYD* 2846 polymorphism T allele for patients on 5-FU monotherapy regimens. Results for this rare variant were not meta-analysed as there was only one patient set per 5-FU-based regimen. However, results were significant for QUASAR2 alone and trending in the same direction for other analyses. All analyses are for the binary grade 012v34 outcome.

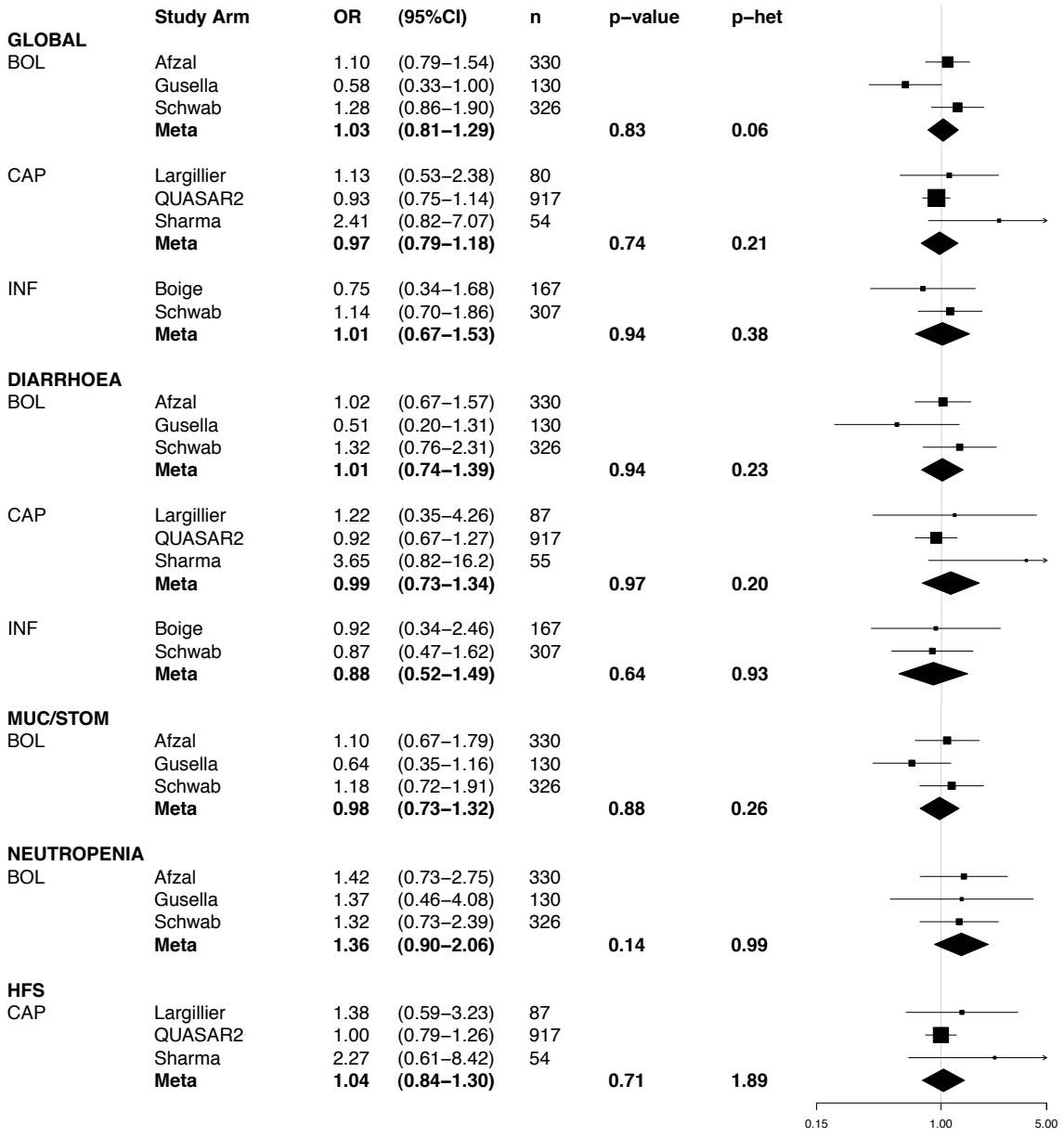


Figure 3.8: Forest plot showing results from the individual patient sets included in the meta-analyses of the *MTHFR* 1298 polymorphism C allele for patients on 5-FU monotherapy regimens. Previously, this variant had been reported as significant, specifically as part of a haplotype with the *TYMS* 3'UTR polymorphism (2) and this variant is included in multiple commercial diagnostic kits. When tested alone and using our standardised global outcome, however, the variant is consistently flat across all studies.

In addition to single-variant analyses, we performed multi-variant analyses for *DPYD* and *TYMS*, as shown in Table 3.12. For capecitabine patients, among others, the *DPYD* *2A and 2846 variants were identified as significant when tested jointly for the presence of either variant versus no variant and the *TYMS* 5'VNTR and 3'UTR were identified as significant when tested in an allelic model for the number of harmful alleles.

QUASAR2 Analyses — Capecitabine Analyses

Polymorphism	Toxicity	N	TAF	OR	95% CI	p-value	S	N	OR	95% CI	p-value
											Meta
DPYD combined model	GLOBAL	863	0.005	5.51	1.95-15.5	0.0013					
*2846A or *2A A allele*	DIARRHOEA	863	0.005	2.48	0.81-7.60	0.11					
	HFS	861	0.005	1.76	0.66-4.71	0.26					
											Logistic
TYMS 5'VNTR adjusted for 3'UTR	GLOBAL	474	0.47	1.24	0.93-1.67	0.15	2	602	1.27	0.98-1.64	0.068
2-repeat allele	DIARRHOEA	474	0.47	1.08	0.70-1.67	0.72	2	602	1.11	0.76-1.61	0.59
	HFS	474	0.47	1.26	0.91-1.75	0.17	2	602	1.20	0.90-1.58	0.21
TYMS 3'UTR adjusted for 5'VNTR	GLOBAL	474	0.69	1.56	1.11-2.18	0.010	2	602	1.42	1.06-1.89	0.017
6bp-insertion allele	DIARRHOEA	474	0.69	1.47	0.88-2.45	0.14	2	602	1.19	0.78-1.81	0.43
	HFS	474	0.69	1.37	0.94-1.98	0.10	2	602	1.40	1.02-1.93	0.038
TYMS score test	GLOBAL	474	0.58	1.38	1.16-1.64	0.00031	2	602	1.33	1.15-1.55	0.00018
number of high-risk alleles	DIARRHOEA	474	0.58	1.24	0.96-1.61	0.096	2	602	1.14	0.92-1.42	0.23
	HFS	474	0.58	1.31	1.08-1.59	0.0063	2	602	1.29	1.09-1.52	0.0030

Infusional 5-FU Analyses — Bolus FU Analyses

Polymorphism	Toxicity	S	N	TAF	OR	95% CI	p-value	S	N	TAF	OR	95% CI	p-value
													Meta
DPYD combined model	GLOBAL	1	321	0.005	4.80	1.06-21.8	0.042	1	325	0.008	5.85	1.63-21.0	0.0068
*2846A or *2A A allele*	DIARRHOEA	1	321	0.005	4.30	0.82-22.7	0.085	1	325	0.008	1.10	0.14-9.18	0.90
	MUC/STOM							1	325	0.008	11.0	3.04-39.8	0.00026
	NEUTROPENIA							1	325	0.008	8.58	2.34-31.5	0.0012
													Logistic
TYMS 5'VNTR adjusted for 3'UTR	GLOBAL	3	957	0.46	1.53	1.14-2.04	0.0040	1	325	0.44	1.08	0.74-1.59	0.68
2-repeat allele	DIARRHOEA	3	941	0.46	1.72	1.11-2.67	0.016	1	325	0.44	1.31	0.76-2.28	0.33
TYMS 3'UTR adjusted for 5'VNTR	GLOBAL	3	957	0.67	0.94	0.69-1.28	0.68	1	325	0.68	1.79	1.12-2.88	0.016
6bp-insertion allele	DIARRHOEA	3	941	0.67	0.93	0.58-1.50	0.77	1	325	0.68	1.90	0.92-3.94	0.083
TYMS score test	GLOBAL	3	957	0.56	1.21	1.01-1.45	0.035	1	325	0.56	1.35	1.06-1.71	0.014
number of high-risk alleles	DIARRHOEA	3	941	0.56	1.29	0.98-1.69	0.070	1	325	0.56	1.53	1.07-2.17	0.018

Table 3.12: Results of the *DPYD* rare allele combined/group test and the *TYMS* single-variant logistic regression and score test are shown for the QUASAR2-only analysis and meta analyses in 5-FU monotherapy regimens, as discussed throughout this chapter. By the group and score analyses, the four included polymorphisms (plus additional rare *DPYD* polymorphisms) are all significant for global capecitabine-related toxicity. Results are generally less significant for the other regimens, but less per-patient data were available for analysis, meaning that differences in significance among monotherapy regimens may be due to sample size differences.

3.15 Consideration of Further Pooling Across 5-FU Regimens

For many polymorphisms, the data we collected came from patients on one of multiple 5-FU-based regimens. In some previous studies that included patients on multiple regimens, particularly from head-to-head trials, pharmacogenetic analyses have been performed across all patients, regardless of drug differences (47; 173). Thus, we considered whether we could more formally combine data from patients across regimens in our meta-analysis.

As discussed in Section 3.12, we observed that while patients on the 5-FU monotherapy regimens all showed some effect from the *TYMS* polymorphisms, this effect was not observed for patients given FOLFOX (infusional + oxaliplatin). Lacking data in FOLFOX for the other significant polymorphisms (*DPYD* *2A and 2846), there was no evidence that data could be shared between monotherapy and combination therapy regimens, with modest evidence (from *TYMS*) that such sharing would be inappropriate.

We sought, then, to further test whether we could formally share information across 5-FU monotherapy regimens, such as by meta-analysis of all underlying patient sets or by meta-analysing the meta-analysis outputs from each regimen. Despite some precedent in the literature for pooling data across regimens, we felt that the burden of proof fell to the data sharing option, given that we had observed large differences in incidence of specific toxicity phenotypes (eg, diarrhoea) among regimens - suggesting a possible mechanistic difference - as was shown in Table 3.4. In order to justify information sharing across regimens, we felt it was necessary to observe a general consistency in effect size estimates across 5-FU-based regimens.

As shown in Figure 3.9, to test consistency of effect across regimens, we plotted effect size estimates for global toxicity and for diarrhoea (the only phenotype

that is common to all three monotherapy regimens) for the 11 polymorphisms with single set/meta-analysis results in all three regimens. (Note that the line extending off the top of the "global" graph extends to an effect estimate with beta greater than 3 but with large confidence intervals.) On initial observation, there was no evident consistency. Upon considering subgroups of polymorphisms, however, some evidence presented. In particular, the rare variants associated with DPYD Deficiency Syndrome (*2A and 2846), shown in red, were consistently estimated to increase toxicity; of note, these estimates did not all differ from zero statistically, particularly for diarrhoea. In addition, for global toxicity, the *TYMS* 5'VNTR and 3'UTR increased risk of toxicity modestly. It was also noted that a general downward trend in effect size estimates across all 11 polymorphisms existed when moving from bolus to capecitabine to infusional, possibly suggesting that associations observed in bolus 5-FU are harbingers of smaller (and thus harder to detect) effects present in other monotherapy regimens.

While these observations may suggest that data sharing would be possible across monotherapy regimens for some polymorphisms, evidence was limited to two classes of polymorphisms: *DPYD* rare variants and the *TYMS* regulatory variants. By nature of the analysis, it was hard to observe consistency in effect among the non-significant variants, and as no additional, individually non-significant variants were included in these two classes (the exception being the *TYMS* G/C SNP, whose effect was inconsistent across monotherapy regimens), we did not feel it was justified to further combine data across regimens for any of the non-significant polymorphsims. In addition, we felt that the inconsistency of effect for diarrhoea merited particular caution; however, it is noteworthy that formal information sharing for the mentioned *DPYD* and *TYMS* variants would improve statistical significance of these biomarkers and allow their effect to be generalised to all monotherapy regimens.

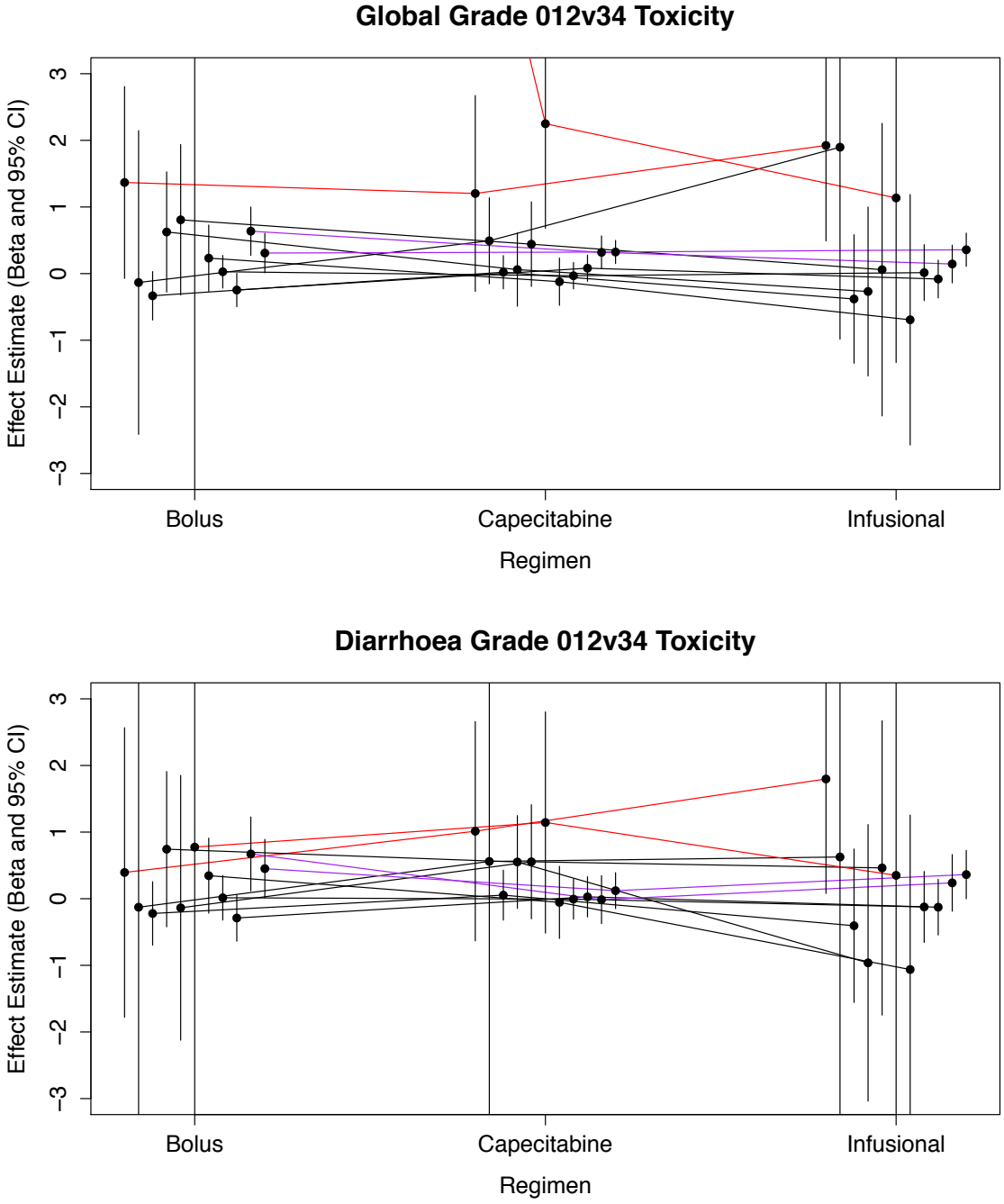


Figure 3.9: Toxicity effect size estimates are compared for consistency across the three 5-FU-based monotherapy regimens, using 11 polymorphisms with data from at least one patient set in each drug regimen. Red lines connect the rare DPYD Deficiency Syndrome variants *2A and 2846. Purple lines connect the *TYMS* 5'VNTR and 3'UTR polymorphisms. Black lines connect the other polymorphisms (all from *DPYD* or *TYMS*). For global, no polymorphism has an effect that is statistically different from zero (no effect) in all three regimens, but it appears that the rare *DPYD* SNPs impact toxicity risk consistently across monotherapy regimens and that the same might be true for *TYMS*. Effect estimates are less consistent for diarrhoea.

3.16 Review of Commercially Sold 5-FU Toxicity Predictive Kits and ROC Analysis in QUASAR2

Three companion diagnostic kits are available commercially for the prediction of 5-FU-related toxicity. None of these kits provide estimates of their predictive power, nor do most justify the inclusion of all selected polymorphisms with citations to the literature. As reviewed in Table 3.13, these kits have varying predictive ability in QUASAR2, but none achieve both high sensitivity and high specificity.

Based on the results of the meta-analysis, at least for capecitabine, the rare DPYD Deficiency Syndrome variants (eg, *2A and 2846) and the *TYMS* 5'VNTR and 3'UTR variants are justified for inclusion in commercial kits for the prediction of toxicity. Our preferred method for inclusion of these variants is as a score test of the number of harmful *TYMS* alleles and a group test counting any rare *DPYD* allele with evidence in Deficiency Syndrome patients, each weighted by the beta coefficient for the variant's impact on toxicity. Applying the effect estimates from the second largest capecitabine study (31), the predictive ability of this test was modeled in 427 QUASAR2 patients (the number genotyped for all tested variants). The resulting receiving operator characteristics (ROC) curve, shown in Figure 3.10, provides an area under curve (AUC) of 0.62 (95% CI 0.57-0.67). At the cut-point at which the maximum proportion of cases was correctly classified (alleles times beta score of 0.70; 64% correctly classified), 6 patients developed high toxicity of the 7 patients predicted to develop it, while 151 additional patients developed high toxicity of the 420 patients predicted to not develop it. This corresponds to a sensitivity of 3.8% (95% CI 1.6-8.5%), specificity of 99.6% (95% CI 97-100%), positive predictive value (PPV) of 86% (95% CI 42-99%), and negative predictive value (NPV) of 64% (95% CI 59-69%) - in total, meaning that we can tell about 1.5% of patients (7 testing positive of 427 total) that they will develop toxicity with 85% certainty, a reflection mostly of the rare *DPYD* variants.

	Test A with 95% CIs	**Test B with 95% CIs**	**Test C with 95% CIs**
True positives	45	156	157
True negatives	218	1	0
False positives	52	269	270
False negatives	112	1	0
Sensitivity p(predicted HT | true increased risk of HT)	29% (22-37%)	99% (95-100%)	100% (97-100%)
Specificity p(predicted LT | normal risk of HT)	81% (75-85%)	0.4% (0-2%)	0% (0-1%)
Positive predictive value	46% (36-57%)	37% (32-42%)	37% (32-42%)
Negative predictive value	66% (61-71%)	50% (3-97%)	Insufficient Data
"Toxicity variants" with QUASAR2 data available	*DPYD* *2A or 2846 heterozygote; *TYMS* 5'VNTR 2R/2R	*DPYD* *2A, 1601(*4), 496 (166), 2846 (949)	*DPYD* *2A, 85(*9A), 2846; *MTHFR* 677, 1298; *TYMS* 3'UTR
"Toxicity variants" with no QUASAR2 data available	Any other *DPYD* variant identified by sequencing of gene (exon and introns)	*DPYD* 1679(*13), 1897(*3), 295-298del(*7), 703(*8), 2983(*10), 62(*12), A551T	*DPYD* 1679(*13), -1590
Interpretation suggested in literature accompanying test kit	High-risk=*DPYD* *2A, 2846A, Ile560Ser or other variants with "significant" effects on protein production or function Moderate risk = 5'VNTR 2R/2R with absence of high-risk variants Low-risk = no *DPYD* variants that have "significant" predicted effect on protein production or function and 5'VNTR 2R/3R or 3R/3R. Uncertain = *DPYD* variants with unproven effects on protein production or function N.B. "Significant effect" is not defined.	No interpretation specified	1) *DPYD* or *TYMS* variant predicts increased sensitivity to 5-FU and may lead to increased risk for 5-FU toxicity 2) Detection of an *MTHFR* mutation is associated with decreased *MTHFR* activity, but impact on toxicity is unclear

Table 3.13: Three commercially available kits are available for the prediction of 5-FU-related toxicity. These kits vary in their included biomarkers and thus their sensitivity and specificity when modeled in the QUASAR2 study. None specify the 5-FU based regimen(s) for which they are valid.

At the 0.248 cut point, sensitivity was 58%, specificity 63%, PPV was 47% and NPV was 75%, with 61% of patients correctly classified. This first cut point functions similarly to commercial Test B while the second functions similarly to commercial Test A. If Test A's sequencing of *DPYD* coding regions - which is not described by the company - functions as a search for rare DPYD Deficiency Syndrome-related variants and excludes the more common synonymous and non-synonymous variants, this test is the strongest available when appropriate allele weightings are applied.

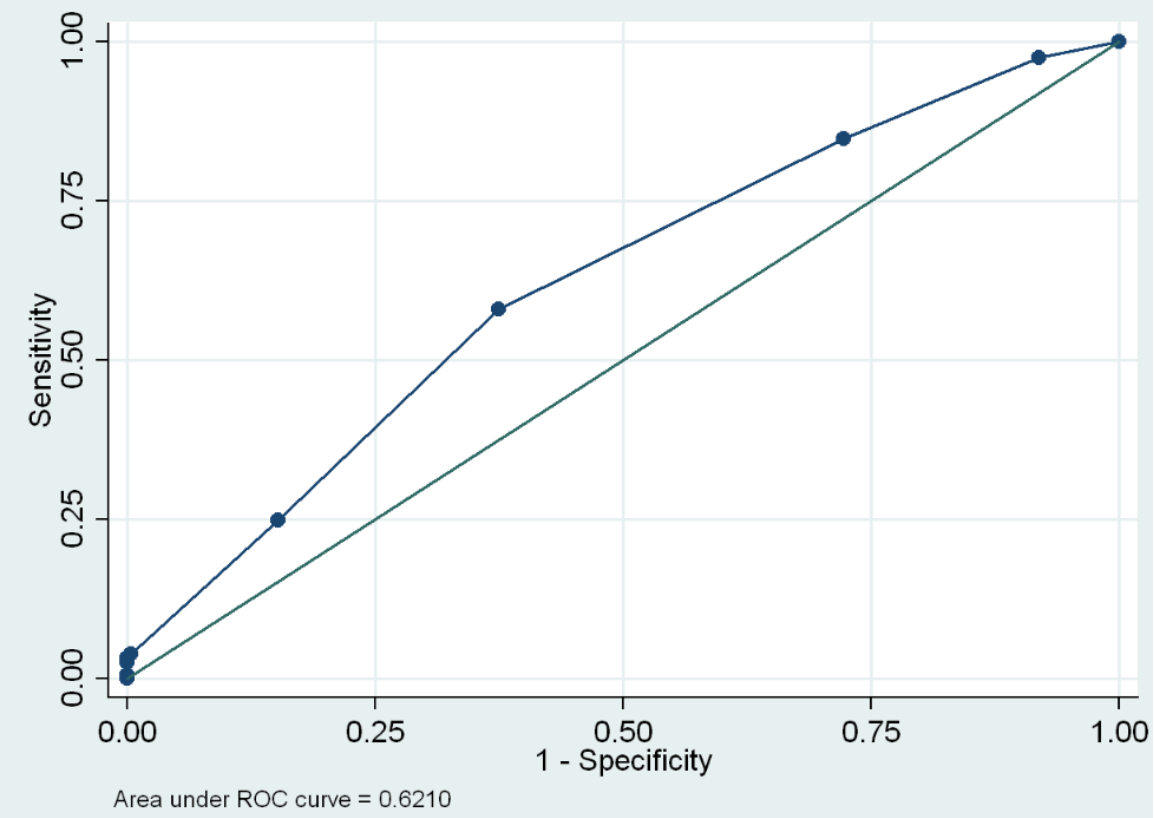


Figure 3.10: Receiving operator characteristics (ROC) analysis was performed in QUASAR2 for the two *TYMS* and two *DPYD* variants identified by the meta-analysis as significant for predicting capecitabine-related toxicity.

3.17 Discussion

Through analysis of 4,855 patients in QUASAR2 and 16 published studies, four alleles have been identified as significantly impacting capecitabine-related toxicity: the two

common *TYMS* UTR variant alleles 5'VNTR 2R and 3'UTR 6bp ins and the two rare *DPYD* variant alleles believed to cause DPYD Deficiency Syndrome, 2846 A and *2A A. In addition, strong evidence exists for rare *DPYD* variants in the same Deficiency Syndrome functional class as *2A and 2846, of which 15 have been found in the literature (reviewed in Table 3.10). In addition to providing meta-analysis, our analysis contained approximately twice the patients and four times the variants of previous reviews of 5-FU-related toxicity, namely the August 2010 Blue Cross Blue Shield Technology Evaluation Center report on 5-FU toxicity biomarkers.

The functional evidence for the *TYMS* UTR variants is less strong than for *DPYD*, including two *in vitro* studies that use a luciferase construct to identify USF binding sites (131) and demonstrate that cell lines with endogenous genetic *TYMS* UTR differences have different gene expression (222). These experiments, however, do not definitively show the functional role of the *TYMS* variants, such as has been done through site-directed mutagenesis for DPYD Deficiency Syndrome alleles (149). As the *TYMS* variants are sufficiently common, their effect can be, and has been, identified by association study, but it remains unknown as to whether the variants are functional sites or tags. This relationship is discussed further in Chapter 4.

The preferred means of measuring the effect of these four variants on toxicity is as a score test of the number of harmful *TYMS* alleles and as a group test comparing any DPYD Deficiency Syndrome causal allele against all other alleles at these loci; compared to single-locus analyses, these increase statistical power and capture the most likely genetic mechanism. As presented in Section 3.14 and Section 3.15, the individual variants have a consistent direction of effect among 5-FU monotherapy regimens and as score/group tests, statistical significance is achieved for capecitabine and suggestive significance is achieved for bolus and infusional 5-FU for global toxicity. The evidence remains ambiguous as to whether capecitabine associations for common *DPYD* variants and for variants in other genes (possibly aside from the

TYMS UTR) can be generalised to other 5-FU monotherapy regimens, which differ in incidence of specific toxicity phenotypes; a third rare *DPYD* variant measured for toxicity in multiple regimens would be helpful here. No variants were significant for any combination therapy regimen, including for FOLFOX, for which data from over 650 individuals were available for both *TYMS* variants (though no data were available for the rare *DPYD* variants). The *TYMS* score and *DPYD* group test together provide only modest predictive power, with an area under curve of 0.62 for capecitabine, potentially less for other monotherapy regimens and likely less for combination therapy regimens.

The four associated variants come from an analysis of 36 previously analysed variants, of which 16 have been reported with at least suggestive significance ($p<0.10$) in one or more study, for some phenotype/regimen combination (Table 1.5). No other variant showed strong evidence of an association, even though many tests were statistically powered (Section 3.14 and Appendix). In some instances, such as for the *MTHFR* variants, lack of significance might have been predicted from previous work in chemotherapy response (104; 224), whereas the the positive significance for *TYMS* was suggested previously by meta-analysis (104), though not for specific regimens.

In addition to power for some analyses, our analysis was limited by several factors. Little precedent exists in the literature for this type of analysis. While we sought to adhere to PRISMA guidelines for systematic reviews and meta-analyses, the guidelines are designed for testing a single outcome, whereas it was necessary for us to refine our objectives during the study: we gathered data from the systematic review in order to determine the list of previously studied polymorphisms and whether specific toxicity phenotypes had consistent incidence across regimens, and we determined whether data from different 5-FU regimens could be combined into single analyses after running a first round of analyses on separate 5-FU regimens. Our finding that formal meta-analysis across regimens was not warranted meant that often few patient

sets were combined for the meta-analysis of a variant in a given 5-FU regimen. We subsequently considered other means of information sharing, such as setting a known direction of effect based on results in one regimen so as to adhere to a one-tailed statistical test, but we could not justify extending patterns observed in one functional class (eg, rare *DPYD* variants) to another (eg, common *DPYD* variants). We also considered whether a more involved Bayesian model might appropriately share information, but felt that the results would be difficult to interpret and replicate. By not sharing information, we may have lost statistical power by reduced sample size; conversely, inappropriately combining data sets across regimens might have obfuscated true signals.

The inability to combine data resulted in a larger multiple testing burden, which we corrected for by applying a significance threshold determined by false discovery rate (FDR), a more conservative measure than most analyses (eg, $p<0.01$ for the exploratory study and $p<0.05$ for replication (66)). Our multiple testing burden was further increased by testing both many polymorphisms (largely independent and adjusted for by FDR) and many toxicity phenotypes (somewhat independent and adjusted for by interpretation as "sensitivity analyses").

We were further limited by our inability to incorporate covariates into analyses, due to our use of allelic counts per patient set (not per patient). This explains why some results previously reported in individual studies were not significant in our analysis, for that same study. It is unknown whether by simplifying the analysis we discarded true signals or whether by standardising the analysis we avoided including spurious associations. Our rationale for collecting counts was two-fold: a practical necessity based on the fact that not all authors wish to share per-patient data and an analysis decision based on the assumption that some studies would make adjustments that improve significance in their particular study but that are not generally meaningful (ie, replicable). Preferable might have been to collect from all authors summary

statistics for analyses made by identical methods, though equivalent covariate data were not obtained by all authors.

In conclusion, while we performed a wide-reaching and rigorous analysis of pharmacogenetic data from the past decade and beyond, we remain unable to adequately predict 5-FU-related toxicity in the clinical setting. We have reduced the field from over 50 polymorphisms to four and identified an effect for each in a large capecitabine trial, but would benefit from further validation of these variants in a second large capecitabine patient set. Further markers are needed in order to reach clinical utility.

Chapter 4

Candidate Gene Analysis of Capecitabine-Related Toxicity

4.1 Introduction

4.1.1 Background

As reviewed in Introduction and Chapter 3, few genetic biomarkers of capecitabine-related toxicity have been determined with high confidence. While the drug activation, action and degradation pathway of 25 candidate enzymes is well established (Figure 1.1), polymorphisms for the genes encoding these proteins have only been investigated by association study for five of the included genes (Table 1.5): *CES2*, *CDA*, *TYMS*, *MTHFR*, and *DPYD*. Within these genes, mostly common coding variants have been investigated, and where rarer or less obviously functional variants have been studied, there has usually been some prior rationale (such as evidence of an effect in other diseases or an impact on protein activity). Thus, a less hypothesis-driven investigation into the polymorphisms in and around the capecitabine/5-FU pathway genes is lacking.

Intriguingly, statistically significant associations with toxicity have been identified

for polymorphisms in two of the candidate genes, *DPYD* and *TYMS* (Chapter 3), suggesting that further investigation into these candidate genes and others might prove fruitful.

As DPYD catabolises some 80% of 5-FU prior to it having a cytotoxic effect at the tumour (84), the mechanism by which decreased DPYD activity results in increased 5-FU-related toxicity is clear. For the two rare *DPYD* variants with strong statistical support – *DPYD* *2A and 2846T>A – the functional impact on DPYD activity is also clear: both are rare variants with a demonstrated impact on enzyme activity (149) and an established role in causing DPYD Deficiency Syndrome when held as compound heterozygotes (Table 3.10) (203; 205; 206). The functional effects of the common and intronic/intergenic *DPYD* variants – of which none have been identified with high confidence – are less clear, but would warrant further investigation if a variant with clear effect was identified by association study.

As reviewed in Chapter 1.2.4, TYMS is the rationally decided target of 5-FU, whose catabolite FdUMP binds to TYMS in place of the endogenous dUMP, inhibiting the enzyme and in turn blocking production of the only source of the DNA monomer precursor to thymidine (82; 127; 211). In regards to tumour cytotoxicity, then, we would expect that cells with low levels of functional TYMS are more susceptible to 5-FU-induced death, as less 5-FU drug is necessary to saturate all TYMS and thus block DNA synthesis. This impact – of decreased *TYMS* mRNA levels increasing drug efficacy and decreasing drug resistance – has been demonstrated (117; 120).

The mechanism by which altered *TYMS* expression levels and activity affect 5-FU-related toxicity – and thus the expected impact of functional variants in and around this gene – is less obvious and has been more controversial, particularly when considering the functional impact of *TYMS* polymorphisms. Two *TYMS* polymorphisms, both of which are in upstream/downstream regulatory regions, have strong statistical support for an association with 5-FU-related toxicity: the *TYMS* 5'VNTR 2-

repeat/3-repeat polymorphism and the 3'UTR 6bp insertion/deletion polymorphism. The 5'VNTR polymorphism was identified first and has been studied most. As early as 1995, Horie et al (88) used a transient transfection assay with induced deletions to show that the 5'VNTR affects efficiency of mRNA expression. Subsequently, Pullarkat et al (160) showed that the 3R allele increased mRNA expression, decreased drug response and decreased drug toxicity (as found in Chapter 3), Kawakami et al (112) showed that the 3R allele increased protein expression via increased translational efficiency and Mandola et al (131) showed that the G-allele of the G/C SNP within the 5'VNTR 3R allele further increases transcriptional activity, among others.

The methods used in these studies have subsequently been criticised by Ghosh et al (69), who note that authors either used the Kozak sequence (a stretch of DNA known to affect translational efficiency) of the firefly – which is weaker than that of the human – or inadequately controlled for experimental error through error bars, replicate experiments and normalisation of reporter activity to account for DNA contamination that would distort the baseline mRNA levels used to determine translational efficiency. In their own study, Ghosh et al re-performed previous experiments using modified methods (ie, testing both firefly and human Kozak sequences, using seven plasmid DNA preparations of two replicates each, adjusting for baseline mRNA levels, and applying statistical analysis) and found that protein expression was not statistically different between 2R and 3R genotypes for either the firefly or human construct. The effects of *TYMS* polymorphisms on mRNA levels were not directly refuted.

4.1.2 Aims and Overview

While both strong functional and outcome data exist for rare *DPYD* variants, the literature lacks a comprehensive investigation of tagging SNPs and rare variants in both this gene and the other 24 5-FU candidate genes, and the functional roles of poly-

morphisms flanking *TYMS* remain ambiguous in regards to their impact on gene expression and 5-FU toxicity. In order to broaden our understanding of the impact that genetic polymorphisms in 5-FU/capecitabine pathway genes have on toxicity, first common and rare variants from SNP arrays, imputation and exome arrays were used to identify which candidate genes held either individually significant SNP associations with toxicity or gene-wide SNP set-based increases in significance. Subsequently, *DPYD*, *TYMS* and *UPP2* were sequenced based on these results. A pathway-based analysis was selected in order to reduce multiple testing when considering genes with a strong functional rationale for impacting 5-FU/capecitabine toxicity. Additionally, we were interested in testing whether the pathway as a whole demonstrated increased associations with toxicity even when individual SNP associations might require extra statistical power to detect.

4.1.3 Patient Inclusion

Three sources of patient clinical and genetic data were used in this analysis (Methods, Section 2.1): QUASAR2 patients formed the primary set and replication of associations was performed using the Spanish Capecitabine Pharmacogenetic (SCP) study and the EPICOLON study. From QUASAR2, of the 1,119 patients with bloods collected as of July 2010, 1,046 were included in either the SNP array (n=12) or exome array (n=102) analysis, or both. Clinical characteristics for these 1,046 patients are shown in Table 4.1. 200 of these patients with the highest and lowest levels of capecitabine-related toxicity were further included in the Roche/454 sequencing project. SCP patients were recruited from one of several Spanish oncology units, with preferential recruitment for patients with high (grade 3+) and low (grade 0/1) toxicity. Total recruitment consisted of 233 colorectal and breast cancer patients, treated with either a standard capecitabine regimen (1250 mg/m^2 twice daily d1-14 every 3 weeks) or a continuous capecitabine regimen (800 mg/m^2 orally every

12 hours daily). EPICOLON patients receive various 5-FU-based regimens (1). 85 patients from this trial were included, all of whom received capecitabine monotherapy.

4.1.4 Toxicity Data

All toxicity phenotypes were collected according to the NCI Common Toxicity Criteria for Adverse Events (CTCAE) version 3.0. Toxicity data scoring was the same as described in Chapter 3. Data were collected for the common capecitabine/5-FU-related toxicity phenotypes: diarrhoea, nausea and vomiting, mucositis, stomatitis, neutropaenia, thrombocytopaenia, and HFS. For each phenotype, the highest grade event (0-4) at any cycle of treatment was recorded as a patient's score for that phenotype, and for QUASAR2 we subsequently used this score to form a "global" measure of the maximum phenotype score observed at any cycle of treatment. For SCP, only handfoot and diarrhoea data were collected, with active selection for sampling and genotyping of patients experiencing high (grade 3+) and low (grade 0/1) toxicity. For EPICOLON, only diarrhoea was recorded. A summary of the incidence of toxicity events experienced within QUASAR2, SCP and EPICOLON is provided in Table 4.2.

Table 4.1: Clinical characteristics for the 1,046 QUASAR2 patients included in the capecitabine/5-FU candidate gene analysis of germline biomarkers for toxicity.

	No.	%
Site		
Colon	930	89%
Rectum	116	11%
Stage		
II	365	35%
III	681	65%
Ethnicity		
Caucasian	1046	100%
Sex		
Male	593	57%
Female	453	43%
Age, years		
Median	65	
Minimum	22	
Maximum	85	
WHO performance status		
0-1	1046	100%
Treatment		
Capecitabine (cap)	496	47%
Cap + bevacizumab	550	53%
Grade 3+ adverse events		
Global	353	34%
Diarrhoea	109	10%
HFS	247	24%
Mucositis	11	1%
Stomatitis	12	1%
Vomiting	15	1%
Neutropaenia	22	2%
Thrombocytopaenia	4	0%

Table 4.2: Incidence of collected toxicity phenotypes from capecitabine treatment for patients from the QUASAR2, SCP and EPICOLON studies. Of note, rates in the SCP study differ, a reflection of the recruitment of grade 0/1 and grade 3/4 HFS and diarrhoea patients.

		QUASAR2		SCP		EPICOLON	
Adverse Event	Highest CTCAE grade reported during treatment	Patients	%	Patients	%	Patients	%
Global	0	75	7%				
	1	241	23%				
	2	375	36%				
	3	334	32%				
	4	19	2%				
	Unreported	2	0%				
Diarrhoea	0	370	35%	143	61%	50	59%
	1	388	37%	23	10%	18	21%
	2	175	17%	17	7%	13	15%
	3	99	9%	32	14%	3	4%
	4	10	1%	4	2%	1	1%
	Unreported	4	0%	14	6%	0	0%
Handfoot	0	176	17%	85	36%		
	1	287	27%	28	12%		
	2	331	32%	30	13%		
	3	246	24%	86	37%		
	4	1	0%	2	1%		
	Unreported	5	0%	2	1%		
Mucositis	0	734	70%				
	1	248	24%				
	2	49	5%				
	3	11	1%				
	4	0	0%				
	Unreported	4	0%				
Stomatitis	0	718	69%				
	1	244	23%				
	2	67	6%				
	3	11	1%				
	4	1	0%				
	Unreported	5	0%				
Vomiting	0	817	78%				
	1	134	13%				
	2	74	7%				
	3	12	1%				
	4	3	0%				
	Unreported	6	1%				
Neutropaenia	0	921	88%				
	1	71	7%				
	2	28	3%				
	3	17	2%				
	4	5	0%				
	Unreported	4	0%				
Thrombocytopaenia	0	961	92%				
	1	67	6%				
	2	9	1%				
	3	0	0%				
	4	4	0%				
	Unreported	5	0%				

4.2 Imputation of SNP Array Data – Strategy, Method and Technical Results

SNP array data included in this analysis were generated previously as part of prior studies (60). In summary, after Dr Palles performed genetic quality control and I reviewed the clinical data for completeness, SNP genotypes were available for 940 patients genotyped on one of three SNP array platforms: Illumina HumanHap300 duo- or HumanHap370 duo- (484 patients), HumanHap610 quad- (364 patients) and HumanOmni2.5 8-bead (92 patients) chips. I performed imputation of these data jointly with Dr Palles. We followed standard imputation methods described in Section 2.4 of Methods and, as described below, developed a patient pooling and downstream filtering strategy that addressed a chief concern: that pooling of imputed genotypes from patients genotyped on multiple different SNP array platforms would cause inflation of results. (For example, case/control allele frequencies could be distorted when imputation of a test SNP was biased toward different reference set alleles based on available direct genotype data on different arrays.) While this potential bias concerned us, we faced less potential results inflation than if our cases and controls were genotyped on separate arrays – our case/control (ie, global grade 012v34 high toxicity vs low toxicity) ratio was approximately 1:3 on each platform.

Prior to performing imputation genome-wide (a computationally burdensome and somewhat slow process), I calibrated the imputation strategy. To do so, I compared the inflation of results upon imputation from a tractable piece of the genome with the inherent inflation to the data (from directly genotyped SNPs), as was the emailed suggestion of Dr Bryan Howie (imputation co-developer). Described below is the analysis that I performed on data from chromosome 18 alone in order to select our patient/array pooling and filtering strategies. Specific methods for implementing this strategy genome-wide are described in Methods (2.4 and 2.4), for which Dr Palles

performed the phasing and imputation according to our method, while I performed association testing by SNPTEST and downstream filtering. The method described here applies to the genotype data in both this chapter and in Chapter 5.

4.2.1 Array Pooling Strategy Options

As per Dr Howie's suggestions, three strategies existed for how to maximise the number of imputed genotypes of patients from multiple arrays without inflating results due to spurious imputation/associations:

1. Impute using only those SNPs present on all three arrays (in this case about 200k SNPs, whereas the largest array has 2,500k).

2. Impute the subjects typed on each array separately, using all of the SNPs on each array (300k, 600k, 2,500k).

3. Impute the subjects typed on each array from a cumulative reference panel of all denser arrays (ie, [1000G] -> [Omni 2.5], [1000G+Omni 2.5] -> [Hap610], [1000G+Omni 2.5+Hap610] -> [Hap370]).

The first option has the least potential for inflation but uses the fewest input SNPs when imputing, minimising the power of this method. Alternatively, the second and third option use all available SNPs on each array as input. The second option is computationally and manually simpler than the third, but is somewhat more prone to cause spurious inflation than the third option, particularly when cases and controls are genotyped on separate arrays. Given that high toxicity and low toxicity patients (ie, cases and controls) were genotyped in equal ratios on each array, I first tested the inflation that resulted from Strategy 2, with the plan of testing Strategy 3 if too much inflation was observed. I did this using chromosome 18, given that it contained a known (ie, directly genotyped) association with toxicity that could act as a positive control.

4.2.2 SNPTEST Downstream Filtering

Prior to testing inflation of the test statistic via Strategy 2, it was necessary to identify appropriate downstream filtering methods to exclude from the association testing output (ie, SNPTEST2 output) those variants that impute inaccurately on average across the three arrays (as measured by the SNPTEST info score). It was necessary to apply these filters to the chromosome 18 test results, otherwise the imputation method being tested might be falsely identified as inherently inflated when it merely required downstream filtering. In addition to deviation from Hardy-Weinberg equilibrium (filtered at the end), two SNPTEST downstream filters were considered: MAF and SNPTEST2 info score. To choose the optimal filtering thresholds, the exclusion criteria were calibrated by applying increasingly stringent levels of filtering until the chromosome 18 positive control peak of the manhattan plot was eroded (results in Table 4.3). Above this level, we were wary of false negative findings, whereas below this level we were wary of false positive findings, though the calibration is imperfect because imputed genotype associations are not equally resilient to strict quality control (eg, due to modest vs common MAF). We see that at MAF below 1%, lambda is higher than when including only common SNPs (pattern not present for directly genotyped SNPs because all were MAF>1%). In the absence of an increased likelihood that rare SNPs impact toxicity, this suggests that rare SNPs either impute less accurately than do common variants or that they are more prone to chance findings, suggesting that rare imputed SNPs should not be tested. Using this allele frequency threshold, we then observed that increasingly stringent SNPTEST2 info scores up to 0.95 further reduced lambda and removed the presence of a modest p-value peak of modest MAF only present from imputed SNPs, with minimal erosion to the positive control peak (including its rare imputed SNPs). Thus, the downstream filters were determined as MAF>1% and SNPTEST2 info scores>0.95.

Filtering	SNPs	Lambda	QQPlot Departure From Expected P-value	Number of Associations at P<1e-03	Note
Direct, no filtering	9902	1.05	~1x10^-3	15 (0.15%)	Baseline, with little evidence of inflation
Direct, MAF>1%	9902	1.05	~1x10^-3	15 (0.15%)	All Hap370 SNPs have MAF>1%
Imputed, no filtering	336829	1.07	~5x10^-2	684 (0.20%)	Manhattan plot contains many lone SNPs of high significance
Imputed, MAF>1%	242864	1.03	~1x10^-3	328 (0.14%)	Little inflation and 90% power for large effects
Imputed, MAF>5%	168601	1.03	~1x10^-3	263 (0.16%)	Loss of all rare SNPs
Imputed, MAF>1%, IS>0.80	167889	1.02	~1x10^-3	263 (0.14%)	Minimal positive control peak erosion; gain of new peak
Imputed, MAF>1%, IS>0.90	167889	1.02	~1x10^-3	220 (0.13%)	Minimal positive control peak erosion; gain of new peak
Imputed, MAF>1%, IS>0.95	141069	1.01	~1x10^-3	176 (0.12%)	Minimal positive control peak erosion; removal of gained peak
Imputed, MAF>1%, IS>0.99	76524	0.99	~5x10^-4	75 (0.10%)	Manhattan plot positive control peak eroded

Table 4.3: Imputation downstream filtering with MAF and SNPTEST info score was chosen analytically by comparing inflation of imputation results to inflation inherently present in the data, using 9,902 directly genotyped SNPs on chromosome 18 from the Hap300 array and testing for an association with grade 012v34 handfoot syndrome. Highlighted filtering thresholds are the optimal levels when using direct genotyping only, when filtering only on MAF and when filtering on MAF and info score. QQ plots and manhattan plots were created for each analysis and inspected visually for changes. MAF=minor allele frequency; IS=SNPTEST2 info score.

4.2.3 SNPTEST Upstream Filtering

While downstream filtering avoids the association testing of SNPs that do not perform well across arrays, it is possible that SNPs with inaccurate genotyping on a single array would be included when their *pooled* info scores among all three arrays achieved an IMPUTE2 info score of 0.95. Thus, using info scores from individual arrays would be a useful *inclusion* criterion for analysis. One analysis option would be to run SNPTEST2 separately for each array, apply filtering and then meta-analyse the remaining results. This option, however, was not available because only 96 patients were genotyped on the Omni2.5 array, and SNPTEST2 requires a minimum of 100 patients. Instead, we considered the use of the IMPUTE2 info score as an inclusion criteria, as a replacement for the SNPTEST2 info score. It is unclear from the literature whether the SNPTEST2 and IMPUTE2 info scores function similarly. Thus, I compared the sensitivity/specificity of the IMPUTE2 info score, using the SNPTEST2 info score as a gold standard, as shown in Table 4.4. Genome-wide, these scores were highly correlated, with an IMPUTE2 info score threshold of 0.95 demon-

strating 95% sensitivity and 98% specificity to the standardly used SNPTEST info score threshold of 0.95. As such, we used an IMPUTE2 info score threshold of 0.95 as an inclusion criteria for SNPTEST association testing.

4.2.4 Pooling Strategy Comparison

Using these thresholds (upstream inclusion criteria of single array IMPUTE2 info score > 0.95, downstream exclusion criteria of MAF$>1\%$ and pooled array SNPTEST2 info score > 0.95), I tested whether pooling Strategy 2 led to spurious associations, as identified by p-value inflation. In three steps, Figure 4.1 highlights the sources of increased power and spurious associations resulting from pooling the directly genotyped and imputed data from all three arrays into a single analysis, using imputed genotypes in addition to direct genotypes but leaving the array results separate, and when both pooling the arrays and using imputation. Table 4.5 quantifies the changes to lambda that result on chromosome 18 from enacting these changes. These calculations were also performed with the positive control peak removed, as this inherently inflates lambda above the expected value. Through this analysis, it was suggested that our three-array pooling method would not lead to significant inflation, and thus we decided to perform our analysis using fully and partially imputed genotypes used in single association tests for patients on all three arrays.

4.2.5 Summary of Imputation Method Used

In summary, imputation was performed separately for each of the three QUASAR2 SNP arrays, using the standard 5Mb sliding window SNP interval and 250kb flanking region. Association testing by SNPTEST2 was performed by pooling patients on all three arrays; an IMPUTE2 info score of 0.95 or greater (for each individual array) was used as an inclusion criteria. Following SNPTEST2, additional downstream filtering was performed, namely excluding SNPs with MAF<0.01 and IMPUTE2 info scores

Filtering	Remaining SNPs	SNPTEST2 Info Score							IMPUTE2 Info Score							Correctly Included (True Positive)	Mistakenly Included (False Positive)	Correctly Filtered (True Negative)	Mistakenly Filtered (False Negative)	Filtering sensitivity p(included|IS>95)	Filtering specificity p(filtered|IS<95)
		Min	Q1	Median	Q3	Max	Mean		Min	Q1	Median	Q3	Max	Mean							
Unfiltered IMPUTE2 output, Hap300	336,829	0.00	0.71	0.93	0.99	1.00	0.83		0.09	0.70	0.92	0.99	1.00	0.83							
SNPTEST2 Info Score <=0.95 (goal to filter)	186,283	0.00	0.57	0.74	0.87	0.95	0.71		0.09	0.57	0.72	0.86	0.98	0.71							
SNPTEST2 Info Score >0.95 (goal to keep)	150,546	0.95	0.98	0.99	1.00	1.00	0.99		0.52	0.98	0.99	1.00	1.00	0.98		150,546	0	186,283	0	100%	100%
IMPUTE2 Info Score >0.80	218,815	0.34	0.93	0.98	1.00	1.00	0.95		0.80	0.93	0.98	1.00	1.00	0.95		150,251	68,564	117,719	295	100%	63%
IMPUTE2 Info Score >0.90	179,834	0.68	0.96	0.99	1.00	1.00	0.98		0.90	0.96	0.99	1.00	1.00	0.98		148,842	30,992	155,291	1,704	99%	83%
IMPUTE2 Info Score >0.95	147,064	0.85	0.98	0.99	1.00	1.00	0.99		0.95	0.98	0.99	1.00	1.00	0.99		143,181	3,883	182,400	7,365	95%	98%
IMPUTE2 Info Score >0.99	77,257	0.97	0.99	1.00	1.00	1.00	1.00		0.99	0.99	1.00	1.00	1.00	1.00		77,257	0	186,283	73,289	51%	100%
Unfiltered IMPUTE2 output, Hap300, MAF>5%	168,601	0.26	0.94	0.98	1.00	1.00	0.94		0.26	0.94	0.98	1.00	1.00	0.94							
SNPTEST2 Info Score <=0.95 (goal to filter)	48,887	0.26	0.73	0.86	0.92	0.95	0.81		0.26	0.73	0.86	0.92	0.97	0.81							
SNPTEST2 Info Score >0.95 (goal to keep)	119,714	0.95	0.98	0.99	1.00	1.00	0.99		0.91	0.98	0.99	1.00	1.00	0.99		119,714	0	217,115	0	100%	100%
IMPUTE2 Info Score >0.80	151,330	0.73	0.96	0.99	1.00	1.00	0.97		0.80	0.96	0.99	1.00	1.00	0.97		119,714	31,616	17,271	0	100%	35%
IMPUTE2 Info Score >0.90	137,184	0.85	0.97	0.99	1.00	1.00	0.98		0.90	0.97	0.99	1.00	1.00	0.98		119,714	17,470	31,417	0	100%	64%
IMPUTE2 Info Score >0.95	119,705	0.92	0.98	0.99	1.00	1.00	0.99		0.95	0.98	0.99	1.00	1.00	0.99		117,960	1,745	47,142	1,754	99%	96%
IMPUTE2 Info Score >0.99	67,012	0.97	0.99	1.00	1.00	1.00	1.00		0.99	0.99	1.00	1.00	1.00	1.00		67,012	0	48,887	52,702	56%	100%

Table 4.4: In order to use IMPUTE2 info scores as an inclusion criteria for association testing, its filtering tendencies were compared to those of SNPTEST2, the commonly used measure in most GWAS. Despite a tendency in the literature to apply lower IMPUTE2 info scores than SNPTEST2 info scores, these measures were highly similar, suggesting that IMPUTE2 info scores could be used in place of SNPTEST2 info scores in order to test how well SNPs imputed on *single* arrays. These results held both at low (any) MAF and at MAF>5%. Of note, missingness – a factor that is known to differ in the two measures – was low for all SNPs. MAF=minor allele frequency; IS=IMPUTE2 info score.

Inflation Source	Pooled Array(s)	SNP Inclusion	Patient Inclusion at a Given SNP	Number of SNPs	Lambda: Full Chr18	Positive Control SNPs	Lambda: Chr18, positive control SNPs removed
Baseline	Hap300 Only	Intersect of genotyped SNPs on all arrays	Patients genotyped for the SNP	7,010	1.043	26	0.992
Baseline	Hap610 Only	Intersect of genotyped SNPs on all arrays	Patients genotyped for the SNP	7,010	1.031	26	1.025
Pooling of array data	Hap300+Hap610+Omni2.5	Intersect of genotyped SNPs on all arrays	Patients genotyped for the SNP	7,010	1.086	26	1.029
Baseline	Hap300+Hap610+Omni2.5	Union of genotyped SNPs on any any array	Patients genotyped for the SNP	16,542	1.090	50	1.055
Imputation fill-in genotyping	Hap300+Hap610+Omni2.5	Union of genotyped SNPs on any any array	Patients genotyped or imputed for the SNP	13,947	1.105	42	1.051
Baseline	Hap300 Only	Genotyped SNPs from each individual array	Patients genotyped for the SNP	10,077	1.056	32	1.002
Baseline	Hap610 Only	Genotyped SNPs from each individual array	Patients genotyped for the SNP	15,767	1.062	56	1.058
Baseline	Omni2.5 Only	Genotyped SNPs from each individual array	Patients genotyped for the SNP	Can't run -- SNPTEST2 requires n patients>100			
Imputation fine-mapping	Hap300 Only	Union of genotyped and imputed SNPs on each array	Patients genotyped or imputed for the SNP	138,857	1.006	359	0.950
Imputation fine-mapping	Hap610 Only	Union of genotyped and imputed SNPs on each array	Patients genotyped or imputed for the SNP	159,296	1.058	419	1.055
All (ie, actual analysis)	Hap300+Hap610+Omni2.5	Union of genotyped and imputed SNPs on each array	Patients genotyped or imputed for the SNP	126,101	1.101	346	1.036

Table 4.5: Despite huge increases to the number of SNPs with genotype information, little inflation is observed from pooling of array data, fill-in genotyping with imputation, fine-map genotyping with imputation, and utilising all three. This was important to understand when determining the imputation method used genome-wide for analyses in this thesis, as it was possible that underlying sources of bias could be overlooked when combining all three potential sources of bias together in the final analysis. Previously described MAF and info score filtering were applied. "Genotyped" refers to SNPs directly genotyped by an array.

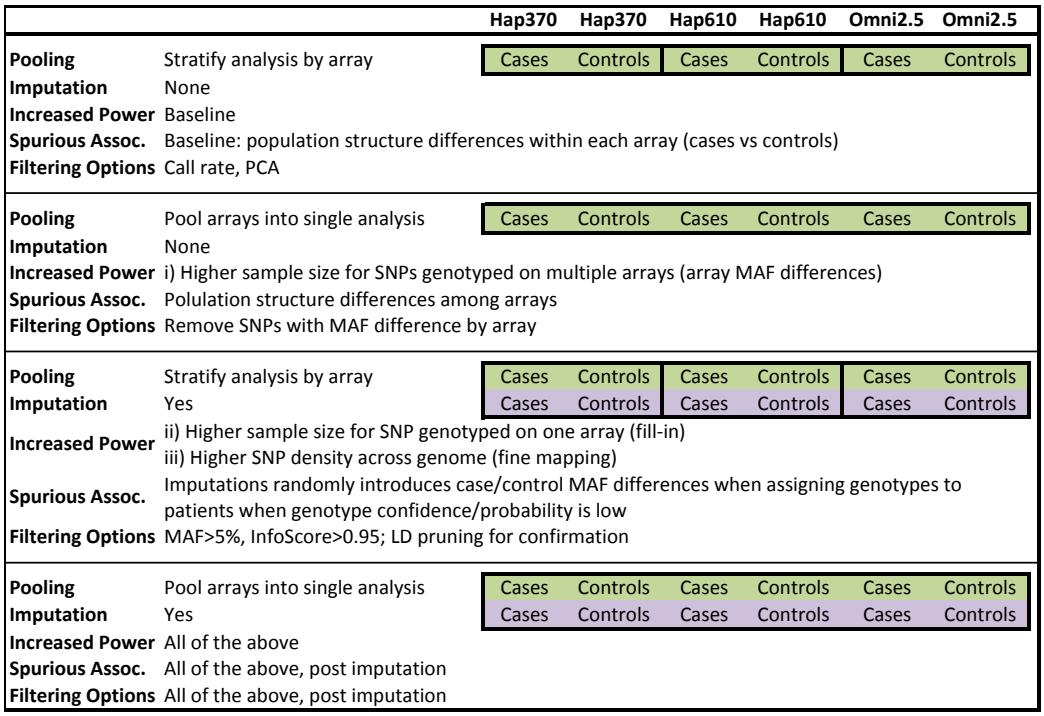


Figure 4.1: QUASAR2 patients were genotyped on one of three SNP array platforms, causing various concerns about forming spurious associations as a result of case/control MAF differences induced or exacerbated by imputation. This figure highlights the strengths (increased power sources) and weaknesses (risk of spurious associations) induced by various patient pooling strategies. The last strategy was used in the final analyses. Green represents directly genotyped SNPs, while purple represents imputed SNPs.

<0.95 (from genotypes pooled across all three arrays), as well as those with a HWE p-value less than 1e-20 for cases and 1e-06 for controls, though no top associations approached these thresholds.

4.3 Exome Array Data Quality Control and Genotype Calling – Technical Results

In addition to the direct and imputed SNP data described in the previous section, the results of this chapter use genetic data from Illumina HumanExome arrays. This section describes the technical results related to the QC and calling of these data, while Section 4.5 describes the association study results.

Using DNA extracted from QUASAR2 and CORGI bloods by Kimberly Howarth, Angela Jones and Dr Claire Palles of the Molecular and Population Genetics group, Oxford, exome array data were newly derived by sequencing centres in Edinburgh and Oxford. I performed quality control (QC) and base calling with zCall of these genotype data, which consisted of 1,234 patients (1,247 samples including duplicates) from the QUASAR2 (n=987) and CORGI studies (n=247), then performed association testing in PLINK for the QUASAR2 patients. Technical results are reported here, while association results are reported later in this chapter and next.

zCall is a series of open-access perl and R scripts that modify standard (eg, Genome Studio) calls, generally made on three genotype clusters, by instead using only the common allele homozygote cluster – the only well populated cluster for rare variants (73). In order to do so, zCall measures the mean and variance of the common homozygote cluster, derived from Genome Studio, and uses it to predict the boundaries of the heterozygote and rare homozygote cluster. This prediction is optimised using common (MAF>5%) variants to determine a "z-score" that scales the cluster boundary lines by a factor of z standard deviations from the mean.

QC measures for exome array data and zCall use were performed according to the guidelines proposed by Dr Anubha Mahajan, Dr Neil Robertson and Dr Will Rayner in the Wellcome Trust Centre for Human Genetics, Oxford (numbers within [ ]s are counts of the patients removed at each step):

1. Perform initial base calling using Illumina Genome Studio.

2. On Genome Studio calls, perform initial removal of obvious per-patient/SNP failures (call rate>90%) [n=3] followed by full per-patient (ie, not per-SNP) QC, including: call rate>98% [n=8], heterozygosity outliers [n=2], gender discordance [n=3], cryptic relatedness (pi-hat>0.1875) [n=4 plus 13 intentional duplicates], and PCA outliers [n=9] (Figure 4.2). At this point, QUASAR2 patients with missing clinical data [n=3] were also removed.

3. Calibrate zCall to determine the optimal z-score, using only the remaining patients/SNPs.

4. Re-perform base calling using zCall, at determined z-score and using only the patients/SNPs passing Genome Studio QC.

5. Using new base calls, re-perform full QC (this time both per-SNP and per-patient), including call rate>98%, heterozygosity outliers [n=1], gender discordance [n=3], GWAS discordance, and PCA outliers.

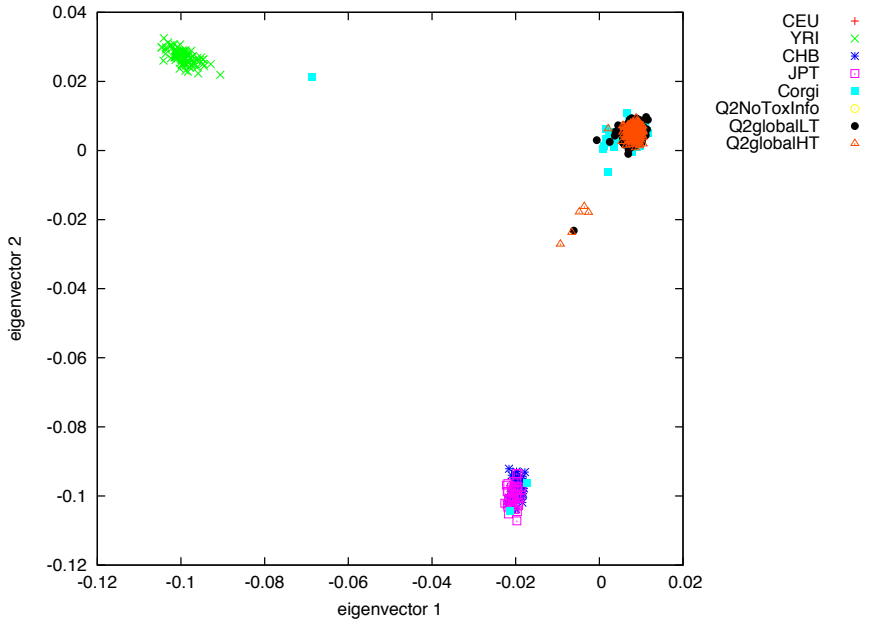


Figure 4.2: Principal components analysis was performed using genetic data derived from common variants on the exome arrays of QUASAR2 and CORGI patients. For QUASAR2, cases (red triangles) and controls (black dots), defined respectively as global 5-FU grade 3/4 and grade 0/1/2 patients, showed strong evidence of similar ancestry. Outliers were excluded from the analysis.

Patients were listed for removal at the end of each analysis and actually removed after either (i) initial Genome Studio patient/SNP QC, (ii) full Genome Studio SNP QC, (iii) clinical data review, or (iv) zCall full QC; as a result, some patient removal counts above include the same patients multiple times. Considering just the 987 included QUASAR2 patients, and counting only the first instance of patient removal,

the following patient removals were made: using Genome Studio calls, 1 for heterozygosity, 6 for gender discordance and 6 as PCA outliers; 3 were removed for clinical data; using zCall calls, 1 for heterozygosity and 2 for gender discordance. This left 968 QUASAR2 patients for analysis.

The optimal z-score was determined as that which maximised the concordance of common SNP (MAF>5%) calls with Genome Studio. Using the open source perl scripts, concordance was evaluated for each of the z-scores from 3 to 13. Concordance ranged from 98.3% (z-score=13) to 99.3% (z-score=6 and 7). Given that a z-score of 7 was found to work best in the original publication, it was selected over a z-score of 6 for this analysis.

zCall markedly altered the base calls made by Genome Studio. GWAS concordance of common SNPs increased from 98.6% for Genome Studio calls, versus the 99.3% noted for zCall. This improvement was seen despite per-patient missingness that fell from a maximum of 10% (excluding patients removed in the initial QC) to a maximum below 1%; as seen in Figure 4.3, missingness fell so greatly that many fewer points are visible on the zCall plots than on the GenomeStudio plots because 936 patients overlap with a missingness below 0.001%. Finally, the placement of the histogram peak for X-chromosome homozygosity (used to identify gender discordance) shifted from a mean around 0.3 using Genome Studio to the expected mean around 0.0 using zCall. These improvements were expected, as they corroborated the findings of the initial zCall method report (73). Thus, calls made by zCall were used for the association studies performed throughout this chapter, filling in genotype calls for rare SNPs that could not be made accurately (and were thus not called) by GenomeStudio. Following QC and calling, 240,339 variants remained for analysis; subsequently, variants were removed if the HWE p-value for cases fell below 1e-20 and 1e-06 for controls, though no top associations approached these thresholds.

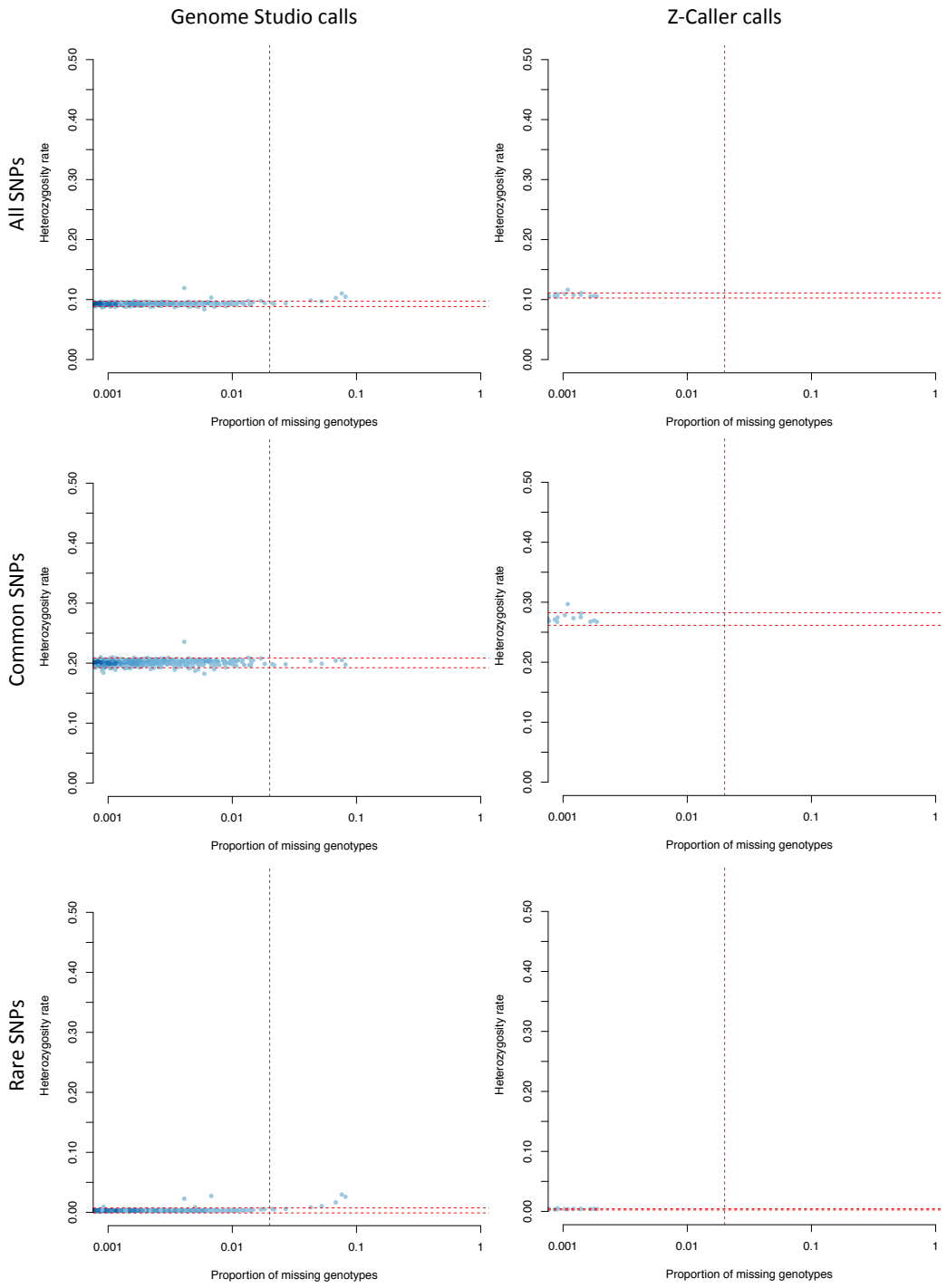


Figure 4.3: The per-patient heterozygosity rate is plotted against the proportion of missing genotypes, using genotype calling from Genome Studio and zCall. When separately considering (i) all SNPs, (ii) only rare SNPs (MAF<0.01) and (iii) only common SNPs, the number of patients with low call rates (ie, below 98%) falls substantially and only a handful of heterozygosity outliers exists. Red vertical line marks the 98% call rate threshold. Red horizontal lines mark 3 standard deviations above and below the mean heterozygosity rate, outside of which outliers were removed.

4.4 Exon Sequencing Analysis – Technical Results

In addition to SNP and exome array variants, the toxicity variant identification results of this chapter use data from next generation sequencing. Methods Section 2.3 fully describes the protocol used for this experiment, while the technical results of this analysis are reported here and the identification of toxicity variants is reported in Section 4.5 of this chapter.

Roche/454 Titanium GS FLX next generation sequencing was used to sequence the exons, splice sites, CpG islands, and UTRs of three genes in the 5-FU/capecitabine pathway (*DPYD*, *TYMS* and *UPP2*), selected based on preliminary (pre-imputation) analysis of GWAS data. The aim of sequencing various 5-FU toxicity candidate gene exons was to search for functional, and likely rare, variants tagged by various SNPs identified within or near to these candidate genes and, secondarily, to find untagged functional variants. Identification of toxicity variants took the form of identifying a pooled frequency difference among 200 patients from the QUASAR2 trial whose DNA was sequenced in one of two groups, a high toxicity ("HT") pool of the 100 patients experiencing the most toxicity (grade 3 or grade 4 diarrhoea in the first 4 cycles of treatment and/or other grade 3/4 toxicities in the first 4 cycles of treatment) and a low toxicity ("LT") pool of the 100 patients experiencing the least toxicity (no events at any cycle of treatment). Promising variants, where not too rare, would be validated in the whole of QUASAR2 and in other patient sets.

Roche/454: Read Depth Achieved

Methods Tables 2.5 and 2.6 show the primers used to amplify the 49 *DPYD*, *TYMS* and *UPP2* target amplicons of this analysis. As described in Section 2.3, the 100 LT patients were combined into 5 pools of DNA from 20 individuals each, as were the 100 HT patients, and thus amplification was performed in 245 (49*5) pools of 20 LT

patients and 245 pools of 20 HT patients. In order to confirm the presence of desired amplicons and the absence of non-desired amplicons and sequencing artifacts, these 490 pools were formed into four LT and four HT pools that could together be used on the 11 wells of one Agilent Bioanalyzer chip. The size (bp) and molarity (pmol/l) of the resulting peaks from each pool were then analysed.

Figure 4.4 shows the results from the first of these eight pools, which contained 16 amplicons of similar length (bp, including 38bp adaptors): 223, 244, 259, 273, 276, 278, 299, 301, 317, 338, 349, 370, 423, 448, 481, and 585. The Bioanalyzer data showed 13 peaks, meaning that not all amplicons were represented by a distinct peak. Peaks representing multiple amplicons are of larger molarity, roughly defined as n-times larger than the molarity seen for most peaks (eg, if most peaks have a molarity around 50pmol/l, then a peak of 100pmol/l likely represents two amplicons because included amplicons are of similar size). 11 of the 16 amplicons in Pool 1 were easily assigned to single distinct peaks with normal molarity and similar size. After assigning these 11 amplicons to peaks, 5 unassigned amplicons and 2 peaks remained. This suggested that in at least one case, multiple amplicons were being captured by a distinct peak. Viewing the molarity measures, the peak at 275bp was treble the normal molarity, suggesting that this distinct peak represented three amplicons of approximately 275bp in length, namely amplicons of length 273bp, 276bp and 278bp. At this point, two amplicons and one peak remained. These two amplicons were of the approximate size of a peak at 463bp with approximately double the normal molarity. Thus, all 16 desired amplicons were present in the Bioanalyzer data for Pool 1. Further, it can be noted that no undesired small (ie, contaminants) or large (ie, non-target amplicons) products existed.

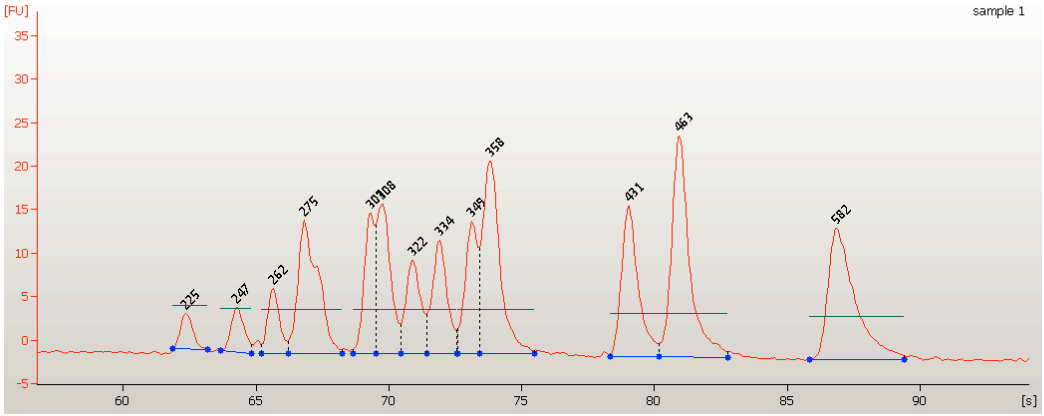


Figure 4.4: An Agilent Bioanalyzer chip was used to confirm the desired PCR amplifications prior to Roche/454 sequencing. One of eight pools is shown here, as described in text.

Roche/454: Quality Control of Amplicon Pools

After all pools were amplified and analysed for quality control, sequencing was performed by the BRC Group of the Wellcome Trust Centre for Human Genetics. As described in Methods (Section 2.3) mapping was performed using Roche GS Reference Mapper software by Dr Kulvinder Kaur in the BRC Group. We aimed to reach 3000-read coverage per 100-patient (HT or LT) pool at each target locus, in order to surpass average 25-read coverage per patient. Initially, 6 amplicons of the LT pool and 3 amplicons of the HT pool achieved below 3,000-read coverage for at least some target loci (Table 4.6).

	Targets with Low Coverage	Amplicon(s) to Re-Sequence	Minimum Reads within Target; Average Reads within Target
Low Toxicity	DPYD E1 + CpG	DPYD E1- part 2	min. = 922; avg. = 5920
	DPYD E26	DYD E26 - part 4	min. = 2453; avg. = 6123
	TYMS E1	TYMS E1- part 1 TYMS E1- part 2	min. = 217; avg. = 2107
	TYMS E5	TYMS E5	min. = 364; avg. = 375
	UPP2 E7	UPP2 E7	min. = 2284; avg. = 2363
High Toxicity	DPYD E1 + CpG	DPYD E1- part 2	min. = 842; avg. = 8686
	TYMS E1	TYMS E1- part 1 TYMS E1- part 2	min. = 365; avg. = 3075
	TYMS E5	TYMS E5	min. = 442; avg. = 455

Table 4.6: Insufficient read depth (below 3,000 reads per locus per 100-patient pool) was achieved for some amplicons after the first round of Roche/454 sequencing.

Amplicons with low coverage were re-sequenced from the 490 PCR pools, having first been checked on a Bioanalyzer. Upon re-sequencing, full coverage was achieved for all exonic target loci except for some in *TYMS* exon 1 and exon 5, as shown in Figure 4.5. In addition, there was low coverage for the UTRs and CpG islands of *DPYD* exon 1 and *TYMS* exon 1 for some loci. The *TYMS* exon 5 amplicon should not be inherently difficult to sequence, given that its G/C content never surpasses 75%; upon re-sequencing, this exon achieved 3,000-read coverage for all LT pool loci and for all HT pool exonic loci except for one exonic locus (!) and the splice site locus of the HT pool, suggesting that possibly this amplicon's concentration was over-estimated prior to inclusion in the HT pool. The low coverage of the *TYMS* and *DPYD* E1 regulatory regions appears to be the result of high G/C content, which reaches as high as 85% and thus might result in hairpin loops or polymerase stuttering resulting from higher thermodynamic stability. Loci not reaching the coverage goal were removed prior to allele frequency comparison between the LT and HT pool.

Roche/454: Variant Filtering

Once sequencing and mapping were completed, variant calling and removal of sequencing artifacts was performed as described in Methods (Section 2.3). Results of this variant filtering protocol are shown in Figure 4.6. Of the 416 variants called by the Roche AVA software, only 39 appeared to not be sequencing artifacts. 31 of these variants were known dbSNPs, while 8 were not listed in dbSNP; as discussed in Section 4.5.9, 18 of these variants had never previously been studied for an association with 5-FU toxicity.

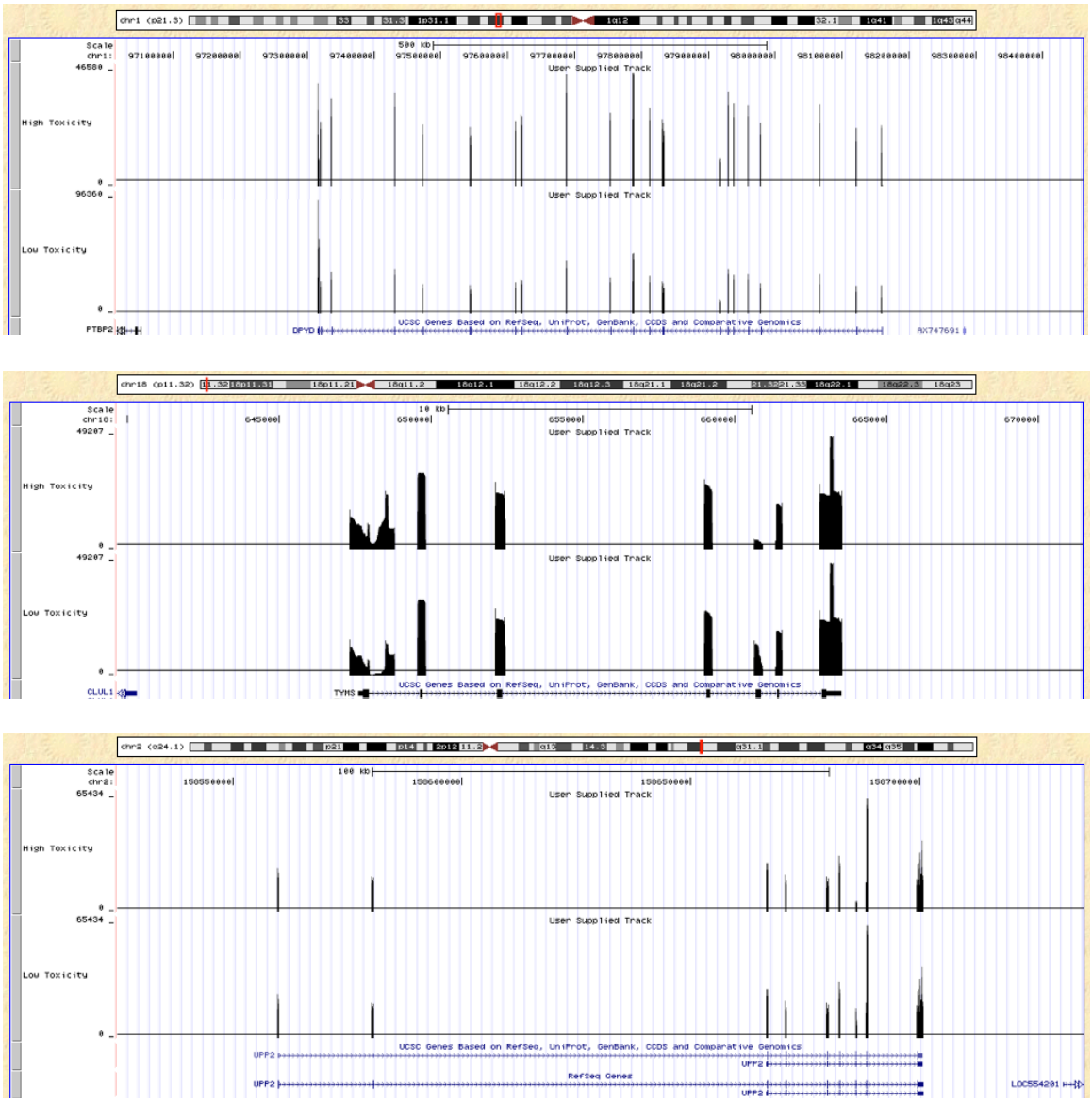


Figure 4.5: UCSC Genome Browser custom tracks were used to visualise the sequencing coverage achieved across all target loci of *DPYD*, *TYMS* and *UPP2* (assessed numerically in Excel). The top panel of each image is the coverage achieved for the 100-patient HT pool, while the bottom panel of each image is for the 100-patient LT pool. The Y-axis range is from 0 coverage to over 40,000 reads per locus per pool, while the horizontal line on each panel marks 3,000 read coverage per pool (ie, 30x coverage per patient) – the coverage target, which was often surpassed over 10-fold.

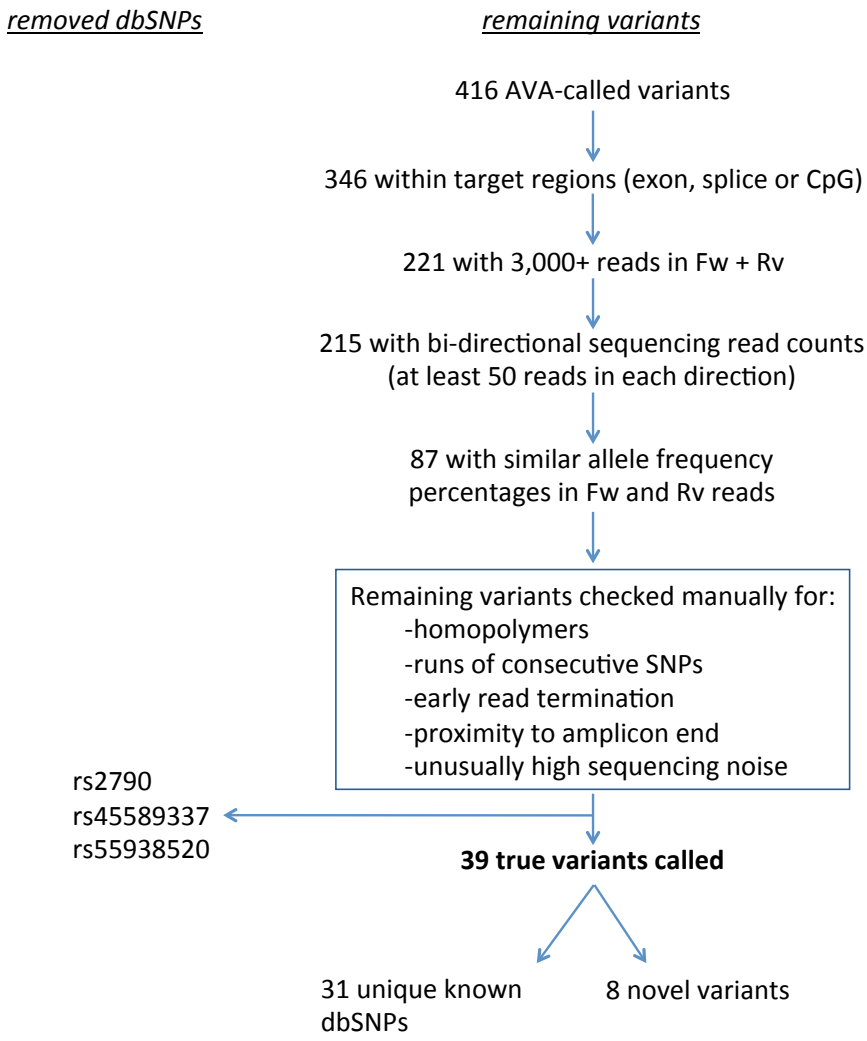


Figure 4.6: Variants identified by Roche/454 sequencing were identified using the AVA software, applying minimal thresholds. A manual filtering scheme (described in Methods, Section 2.3) was then devised and implemented, which reduced the 416 called variants to only 39 true variants. Three dbSNPs were removed in this process, while 31 dbSNPs of varying frequency and sequencing complexity were identified, suggesting that our method could accurately identify most true variants with minimal false positives.

4.5 5-FU/Capecitabine Candidate Gene Analysis: Toxicity Variant Identification Results

4.5.1 Overview of Analysis Method and Included Data

Results of the SNP/exome array and next-generation sequencing analyses of 5-FU pathway gene loci are presented in this section (4.5).

For the SNP/exome array analysis, 1,046 patients were analysed for up to 1,456 genetic variants directly genotyped within 25kb of the 25 5-FU pathway genes – our target region – as shown in Table 4.7. For the SNPs, imputation was performed to genotype missing patients at a locus (ie, "fill-in imputation") when only a subset of patients were genotyped at a locus (due to differences in array content), such that 940 patients were included in the analysis of each SNP that imputed successfully. In hopes of maximising statistical power, biologic relevance and clinical utility, our primary analysis was for associations between global (any 5-FU-related) dose-limiting (grade 012v34) toxicity and genotype. An allelic model was employed by linear and logistic regression, adjusting for age, treatment arm and gender as covariates. SNPTESTv2 and PLINK were used to test SNP associations, while PLINK was used to test exome array associations. A Bonferroni-corrected p-value threshold (3.43e-05) was used to determine statistical significance. Subsequently, regions with significant associations were further investigated by imputing the surrounding region (ie, "fine-mapping") and by testing specific toxicity phenotypes under both binary and quantitative models.

For the sequencing analysis, 100 high toxicity (HT) and 100 low toxicity (LT) patients were compared for an allele frequency difference at 38 loci identified by Roche/454 next-generation sequencing.

Gene Symbol	Gene Name	Location – Build 37 (coordinates do not include 25kb flanking region)	Number of Test Panel SNPs
ABCB1	ATP-binding cassette, sub-family B	chr7:87132948-87342564	77
ABCC3	ATP-binding cassette, sub-family C, member 3	chr17:48712218-48769063	64
ABCC4	ATP-binding cassette, sub-family C, member 4	chr13:95672083-95953687	224
ABCC5	ATP-binding cassette, sub-family C, member 5	chr3:183637724-183735727	101
ABCG2	ATP-binding cassette, sub-family G, member 2	chr4:89011416-89152474	57
CDA	cytidine deaminase	chr1:20915444-20945400	26
CES1	carboxylesterase 1 isoform a precursor	chr16:55836764-55867075	24
CES2	carboxylesterase 2 isoform a precursor	chr16:66968347-66978994	59
DPYD	dihydropyrimidine dehydrogenase	chr1:97543300-98386615	239
DPYS	dihydropyrimidinase	chr8:105391652-105479277	69
MTHFR	methylenetetrahydrofolate reductase	chr1:11845787-11866115	38
PPAT	phosphoribosyl pyrophosphate amidotransferase	chr4:57259529-57301845	29
RRM1	ribonucleoside-diphosphate reductase subunit 1	chr11:4115924-4160106	29
RRM2	ribonucleoside-diphosphate reductase subunit 2	chr2:10262735-10270623	19
SLC22A7	solute carrier family 22 member 7 isoform b	chr6:43265998-43273276	26
SLC29A1	equilibrative nucleoside transporter 1	chr6:44187242-44201888	26
TK1	thymidine kinase 1	chr17:76170160-76183285	35
TYMP	thymidine phosphorylase	chr22:50964182-50968258	92
TYMS	thymidylate synthetase	chr18:657604-673499	34
UCK1	uridine-cytidine kinase 1 isoform a	chr9:134399191-134406655	43
UCK2	uridine-cytidine kinase 2 isoform a	chr1:165796890-165877339	22
UMPS	uridine monophosphate synthase	chr3:124449213-124464040	34
UPB1	beta-ureidopropionase	chr22:24890077-24922553	30
UPP1	uridine phosphorylase 1	chr7:48128355-48148330	16
UPP2	uridine phosphorylase 2	chr2:158851691-158992478	43

Table 4.7: A panel of 1,456 common and rare variants was formed for analysis by extracting polymorphisms within 25kb of the 25 5-FU/capecitabine pathway genes.

4.5.2 Overview of SNP and Exome Array Results

From the SNP and exome array variant analyses, six SNPs achieved the Bonferroni-corrected p-value threshold of p<3.43e-05 with global grade 012v34 toxicity, none of which were from the exome arrays. When testing for independence by logistic regression incorporating multiple SNPs as covariates and minimising the AIC, these reduced to two new SNP associations at *DPYD* (Section 4.5.3 and Section 4.5.4) and one new SNP association at *TYMS* (Section 4.5.5). This third association was subsequently shown to impact the adjacent *ENOSF1* gene, not *TYMS* (Section 4.5.6). All results described here are for the SNP and exome array association tests merged into one results set.

4.5.3 *DPYD* rs12132152 is Associated with Global Capecitabine Toxicity, Driven by HFS

Of the two *DPYD* associations with global capecitabine/5-FU-related grade 012v34 toxicity at p<3.43e-05, the most significant SNP was rs12132152, a G>A SNP located 22kb downstream of *DPYD* at chr1:97,523,004 (build 37/hg19), whose A-allele (frequency of 0.03) was associated with increased risk of 5-FU/capecitabine-related toxicity (OR=3.65, 95% CI=2.07-6.44, p=7.50e-06).

In order to identify the association's functional source (locus/loci) and the toxicity phenotype(s) underlying the global toxicity association, rs12132152 and its surrounding region were tested for associations with global toxicity in a quantitative (01v2v34) model and the specific capecitabine/5-FU-related toxicity phenotypes were tested in both binary (012v34 for HFS/diarrhoea; 0v1234 for rare phenotypes) and quantitative (01v2v34, for HFS/diarrhoea only) models. Both the 1,456 SNP panel and the fully imputed (ie, fine-mapping) SNPs in the region approximately 500kb up- and downstream of *DPYD* were included. None of the rare phenotypes (mucositis, stomatitis,

vomiting, neutropaenia, and thrombocytopaenia) reached even p<1e-03. However, in the *DPYD* region, a Bonferroni-corrected significance level was reached for global, diarrhoea and HFS. Figure 4.7 shows manhattan plots (created in LocusZoom) for these phenotypes, highlighting the LD between SNPs in the region and rs12132152.

Based on the LocusZoom plots, it was apparent that the two *DPYD* associations with global toxicity were independent, given the lack of correlation (measured by r^2) between the two regions, in addition to the lower AIC achieved when including both SNPs in a logistic model for global toxicity, relative to when each SNP was tested alone (see Methods, Section 2.5). Thus, it was apparent that handfoot syndrome, which was significantly associated with rs12132152 and its correlated SNPs – but not diarrhoea, which was associated with SNPs 330kb upstream (discussed in the next section) – was responsible for the association of rs12132152 with global toxicity. It was also apparent from LocusZoom that seven SNPs in very strong LD with rs12132152 – all of which were fully imputed – were slightly more significant than this tag SNP for associations with both global toxicity and HFS (rs76387818, rs150143811, rs116448748, rs75968189, rs12117420, rs12123160, and rs75206280), of which rs76387818 at chr1:97,539,400, with r^2 of 0.98 with rs12132152, was the strongest (OR=3.8, p=4.2e-06 for global grade 012v34).

The associations of rs12132152 with global, diarrhoea and HFS are shown in Table 4.8, in which we observe that the SNP is strongly associated with HFS under both a quantitative and binary measure (as visualised in LocusZoom). The effect on diarrhoea was not significant. The rs12132152 association was independent of the associations identified for other *DPYD* SNPs, both from this study and among the known associations reviewed by meta-analysis in Chapter 3, as identified by LD structure (r^2 of 0 in QUASAR2) and AIC.

The association between global grade 012v34 handfoot syndrome toxicity and rs12132152 genotype was then assessed in 233 patients from the Spanish Capecitabine

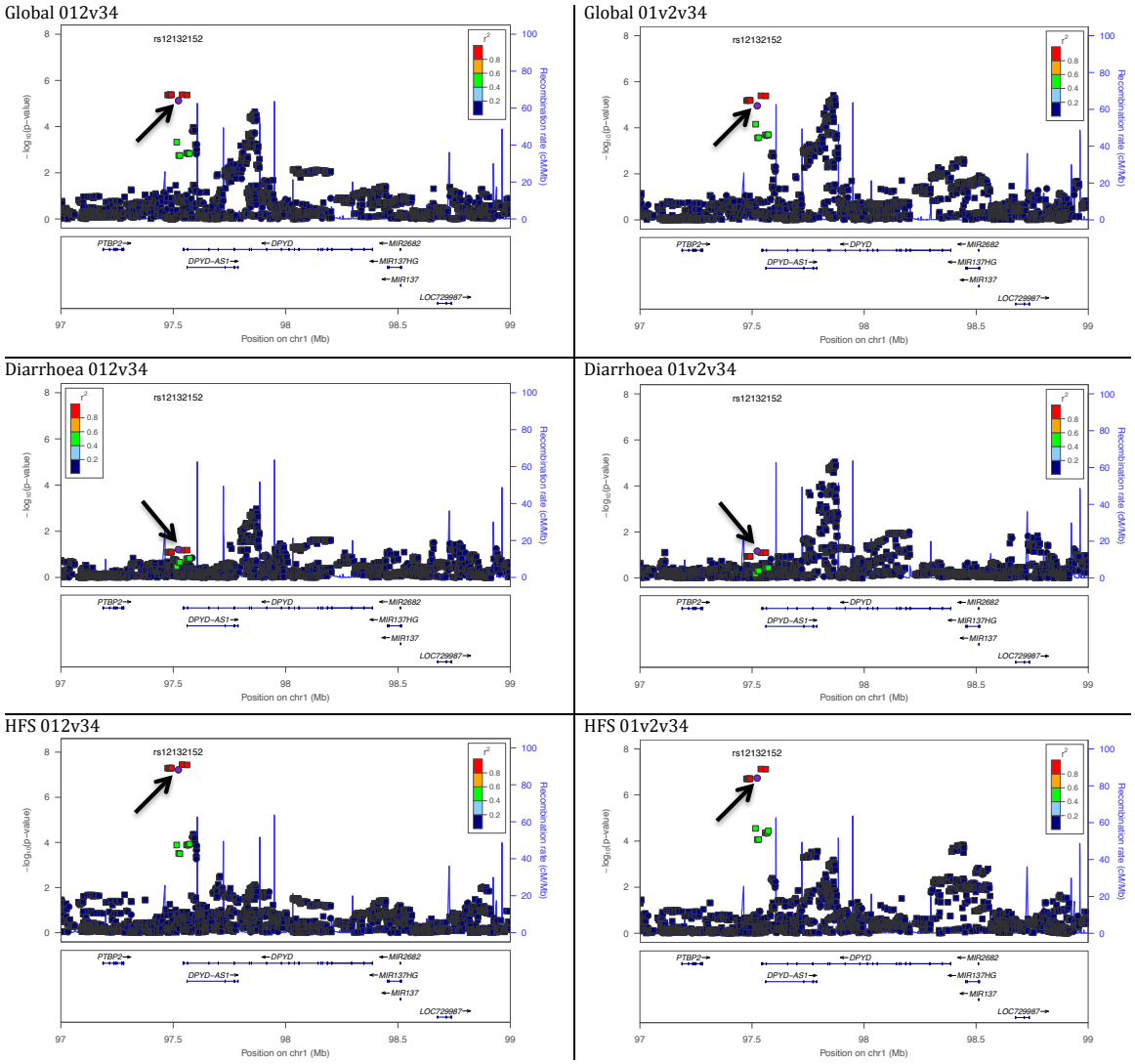


Figure 4.7: LocusZoom plots for associations with capecitabine/5-FU-related toxicity phenotypes for *DPYD* rs12132152 and the surrounding region. The rs12132152 SNP, shown as a purple circle with an arrow pointed at it, is at chr1:97,523,004, 22kb downstream of *DPYD*. Other circles represent associations among SNPs included on the directly genotyped 1,456 SNP/exome variant panel, while squares are fully imputed fine-mapping SNPs. Across the entire region, linkage disequilibrium is displayed by colour, with red SNPs having an r-squared of at least 0.8 with the lead SNP (rs12132152), orange having 0.6 to 0.8, green having 0.4 to 0.6, light blue having 0.2 to 0.4, and dark blue having little LD – all based on the March 2012 release of 1,000 Genomes. Recombination sites are marked by vertical purple lines and the location of *DPYD* is shown below each plot.

Pharmacogenetic (SCP) study. As shown in Table 4.8, contrary to the results observed in QUASAR2, rs12132152 exhibited a protective effect among SCP patients, though this association only achieved modest significance. Possibly this failure to

rs12132152 A-Allele	OR	95% CI	P-value
QUASAR2			
Global Binary: 012v34	3.65	(2.07-6.44)	$7.50\text{x}10^{-6}$
Global Quantitative: 01v2v34	1.59	(1.29-1.96)	$1.11\text{x}10^{-5}$
Diarrhoea Binary: 012v34	0.44	(0.18-1.04)	0.062
Diarrhoea Quantitative: 01v2v34	0.85	(0.71-1.01)	0.068
HFS Binary: 012v34	5.81	(3.07-10.98)	$6.05\text{x}10^{-8}$
HFS Quantitative: 01v2v34	1.74	(1.41-2.14)	$1.85\text{x}10^{-7}$
SCP			
HFS Binary: 012v34	0.30	(0.10-0.86)	0.025

Table 4.8: Associations with common 5-FU/capecitabine-related toxicity phenotypes for *DPYD* rs12132152. No significant associations existed among the rare phenotypes or among other clinically relevant toxicity codings (eg, diarrhoea grade 01v234).

replicate suggests population structure differences among the UK and Spanish patient sets or that the structure of the SCP dataset – selected for high and low toxicity events and with a depletion of patients with modest toxicity – makes the datasets non-comparable, at least for HFS.

The genetic mechanism by which rs12132152 might impact 5-FU/capecitabine toxicity was then investigated using ENCODE data (`http://genome.ucsc.edu/ENCODE/`) and publicly available data for analysis of mRNA expression available from the Genevar and TCGA (The Cancer Genome Atlas) databases. As part of the international ENCODE project, data from scores of functional studies have been compiled onto the UCSC Bioinformatics Genome Browser, and in particular the following tracks were assessed: integrated regulation, chromatin interactions, DNA methylation, open chromatin by DNaseI HS and FAIRE, histone modification, RNA binding proteins, transcription factor binding, replication timing by repli-chip, nucleosome position by MNase-seq, RNA binding protein associated RNA by SwitchGear, and replication timing by repli-seq. The TCGA database contains genome-wide eQTL data (mRNA expression by SNP genotype) directly assessed from tumour and normal colonic tissue. Analysis of TCGA data was performed by Luke Freeman-Mills of the Tomlinson

Lab, according to the methods of Qiyuan et al (121). Three cis eQTL (SNP genotype on nearby mRNA expression) studies exist that have contributed their genome-wide SNP and expression data to the publicly available database Genevar (52; 146; 188), where SNP genotype can be tested for an association with mRNA expression of any gene (cis or trans). The Stranger et al and Dimas et al studies assess lymphoblastoid (for both) and fibroblasts, B-cells and T-cells (for Dimas et al) total RNA extracted from cell lines, while the Nica et al study used punch biopsies for skin and fat cells and peripheral blood to form lymphoblastoid cell lines. The Dimas et al study analyses tissue from 85 individuals' umbilical cords available from the GenCord project, the Stranger et al study analyses 726 individuals available from eight HapMap project populations and the Nica et al study analyses matched tissue samples taken simultaneously from approximately 30 monozygotic and 40 dizygotic twin pairs and 20 singletons.

Using these resources, rs12132152 and its strongly associated SNPs were each assessed. According to ENCODE, rs75968189 and rs12117420 both map to regions of open chromatin, determined by DNase-seq and FAIRE-seq, suggesting that possibly rs12132152 tags the effect of at least one other SNP that binds transcription factors that regulate gene expression. A third correlated SNP, rs12123160, was of particular interest because it lies at a methylated CpG site, where methylation might reduce gene expression. Based on the high 5-FU catabolic activity of *DPYD* in the liver, it would be noteworthy if rs12132152 affected expression of this gene in this tissue; unfortunately, none of the public databases made such data available, and the rs12132152 genotype was not associated with *DPYD* mRNA expression in adipose tissue, lymphoblastoid cells or skin according to the Genevar databse ($p>0.13$) or in colon tissue according to TCGA ($p=0.72$). Thus the specific means by which rs12132152 affects capecitabine/5-FU-related toxicity remains unknown.

4.5.4 *DPYD* rs7548189 is Associated with Global Capecitabine Toxicity, Driven by Multiple Phenotypes

The second *DPYD* SNP associated with global grade 012v34 toxicity was rs7548189, a C>A SNP located intronic to *DPYD* at chr1:97,867,713 (build37/hg19), whose A-allele (frequency of 0.2) was strongly associated (OR=1.67, 95% CI=1.31-2.13, p=3.33e-05).

The same analysis protocol used for rs12132152 was applied to rs7548189 – namely that the SNP's functional source was searched for by fine mapping the region and the causal toxicity phenotype(s) underlying the global measure was investigated by testing phenotypes individually. Figure 4.8 shows the LocusZoom regional association plots provided for rs12132152, but with the linkage disequilibrium colours selected based on genotype correlation with rs7548189. As with rs12132152, none of the rare phenotypes were formally associated with rs7548189 genotype; however, rs7548189 and nearby SNPs were associated with grade 0v1234 vomiting at approximately p=5e-04. Table 4.9 shows the SNP's association with vomiting, as well as with global toxicity and the common 5-FU/capecitabine toxicity phenotypes, among which diarrhoea was formally significant under both the 01v2v34 quantitative model and the clinically dose modifying 01v234 binary model, while HFS was modestly significant under the 012v34 and 01v2v34 models. In the SCP study, rs7548189 showed replication for its association with binary HFS, but did not show evidence of an association with binary diarrhoea. Only data for these two phenotypes were collected.

Unlike rs12132152, it was apparent that rs7548189 might be causal of the "cluster" of multiple toxicity phenotypes characteristic of some high toxicity patients, as mentioned in Introduction. To test whether these phenotypes were independent, or whether associations with some phenotypes were the result of a secondary (confounding) effect from associations between/among phenotypes, rs75481489 was tested in a logistic model for associations with multiple phenotypes (eg, diarrhoea ~ SNP vs

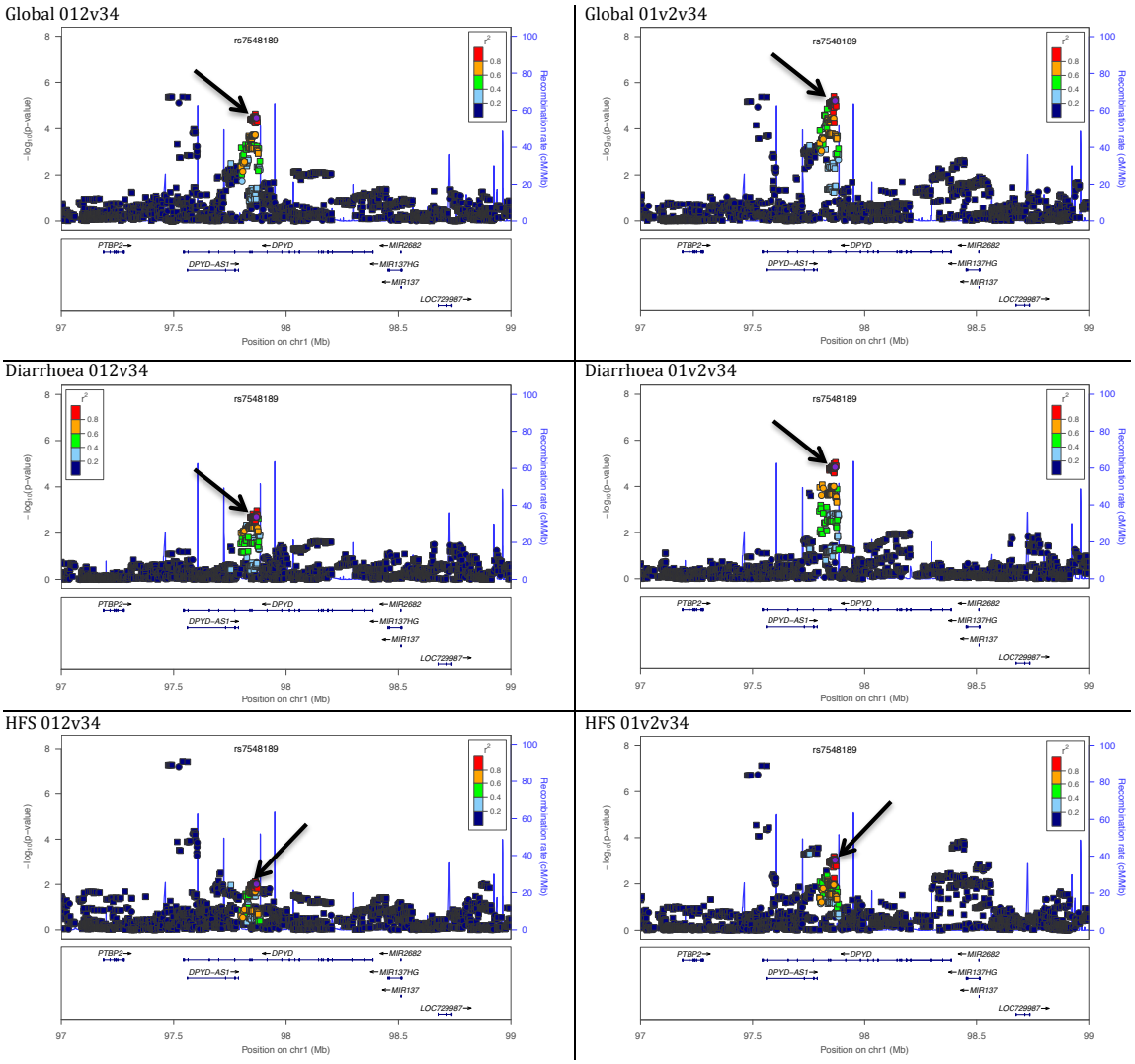


Figure 4.8: LocusZoom plots for associations with capecitabine/5-FU-related toxicity phenotypes for *DPYD* rs7548189 and the surrounding region. The rs7548189 SNP, shown as a purple circle with an arrow pointed at it, is at chr1:97,867,713, intronic to *DPYD*. Other circles represent associations among SNPs included on the directly genotyped 1,456 SNP/exome variant panel, while squares are fully imputed fine-mapping SNPs. Across the entire region, linkage disequilibrium is displayed by colour, with red SNPs having an r-squared of at least 0.8 with the lead SNP (rs7548189), orange having 0.6 to 0.8, green having 0.4 to 0.6, light blue having 0.2 to 0.4, and dark blue having little LD – all based on the March 2012 release of 1,000 Genomes. Recombination sites are marked by vertical purple lines and the location of *DPYD* is shown below each plot.

diarrhoea ~ SNP + HFS vs diarrhoea ~ SNP + HFS + vomiting), as shown in Table 4.10. Considering the global binary measure, it was clear that no single individual phenotype caused the association with global, as modest significance with global tox-

rs7548189 A-Allele	OR	95% CI	P-value
QUASAR2			
Global Binary: 012v34	1.67	(1.31-2.13)	$3.33x10^{-5}$
Global Quantitative: 01v2v34	1.23	(1.12-1.34)	$5.96x10^{-6}$
Diarrhoea Binary: 012v34	1.78	(1.24-2.56)	0.0020
Diarrhoea Binary: 01v234	1.76	(1.37-2.28)	$1.39x10^{-5}$
Diarrhoea Quantitative: 01v2v34	1.18	(1.10-1.27)	$1.46x10^{-5}$
HFS Binary: 012v34	1.43	(1.09-1.88)	0.0097
HFS Quantitative: 01v2v34	1.16	(1.06-1.27)	0.00094
Vomiting Binary: 0v1234	1.64	((1.24-2.17)	0.00056
SCP			
Diarrhoea Binary: 012v34	0.78	(0.39-1.53)	0.46
HFS Binary: 012v34	1.66	(1.00-2.74)	0.048

Table 4.9: Associations with common 5-FU/capecitabine-related toxicity phenotypes for *DPYD* rs7548189, including for all clinically relevant codings according to the QUASAR2 protocol (eg, diarrhoea and HFS grade 01v234), and for rare phenotypes of modest significance.

icity remained upon adjusting for each individual phenotype. Independence of the effect of the SNP on the individual phenotypes was then tested. Using the binary measure of all three phenotypes, the association with HFS appeared independent of the associations between SNP genotype and each of diarrhoea and vomiting, but the associations with each of diarrhoea and vomiting suggested some confounding because they became less significant when adjusting for the other phenotype. Possibly this suggests that the cellular mechanism by which HFS develops differs from the mechanism by which other phenotypes develop. A wider analysis of phenotype independence is provided in the next chapter.

The genetic mechanism by which rs7548189 causes toxicity was investigated using the same resources as for rs12132152. The rs7548189 SNP did not itself show any strong correlation with ENCODE regulatory peaks, but its lead global toxicity fine-mapping SNP – rs12011143, with r^2 of 0.95 – exists in a region of open chromatin with possible weak enhancer activity. rs7548189, and its correlated SNPs as well, was not associated with DPYD mRNA levels in lymphoblasts, fibroblasts, T-cells,

rs7548189 A-Allele	P-value	Beta	AIC
Global ~ SNP	7.89E-05	0.48	
Global ~ SNP + Diarrhoea	3.98E-03	0.40	952.03
Global ~ SNP + HFS	1.20E-03	0.58	565.36
Global ~ SNP + Vomiting	5.61E-04	0.42	1173.5
Diarrhoea ~ SNP	3.13E-03	0.50	638.36
Diarrhoea ~ SNP + HFS	3.00E-03	0.51	640.29
Diarrhoea ~ SNP + Vomiting	2.15E-02	0.41	611.16
Diarrhoea ~ SNP + HFS + Vomiting	2.06E-02	0.41	613.09
HFS ~ SNP	1.40E-02	0.33	1013.8
HFS ~ SNP + Diarrhoea	1.34E-02	0.33	1015.7
HFS ~ SNP + Vomiting	1.36E-02	0.33	1015.0
HFS ~ SNP + Diarrhoea + Vomiting	1.29E-02	0.33	1016.9
Vomiting ~ SNP	5.98E-04	0.47	949.21
Vomiting ~ SNP + Diarrhoea	3.17E-03	0.41	921.84
Vomiting ~ SNP + HFS	6.69E-04	0.47	951.15
Vomiting ~ SNP + Diarrhoea + HFS	3.57E-03	0.41	923.71

Table 4.10: The rs7548189 SNP is significantly associated with binary (012v34) global toxicity and moderately associated with binary diarrhoea (012v34), HFS (012v34) and vomiting (0v1234). Independence of associations between genotype and SNP was tested using logistic models for the binary outcome measure, adjusted for the binary levels of the other toxicity phenotypes. Results differ slightly from Table 4.9 due to the use of rounded integer genotypes from imputation probabilities. AIC within each set of tests was used to identify whether inclusion of additional phenotype data improved the ability of the model to explain the outcome phenotype being tested. All results are for the effect of the SNP on the first listed (outcome variable) phenotype.

adipose tissue, or skin on the Genevar database ($p>0.07$) or in colon tissue from TCGA ($p=0.97$). As with rs12132152, rs7548189 was independent of associations with toxicity of other known variants.

As with rs12132152, rs7548189 was independent of the associations identified for other *DPYD* SNPs (r^2 of 0 in QUASAR2). Among the other SNP and exome variants tested in *DPYD* and its 25kb flanking region, no additional independent variant reached formal significance.

4.5.5 *TYMS/ENOSF1* rs2612091 is Associated with Global Capecitabine Toxicity, Driven by HFS

The same method used to investigate *DPYD* SNPs was extended to the other 24 candidate genes, of which only a *TYMS* polymorphism showed an association with global grade 012v34 toxicity that met the significance threshold. Specifically, the G-allele of rs2612091 (frequency of 0.45) – located at chr18:683,607, 10kb downstream of *TYMS* within an intron of the adjacent/overlapping *ENOSF1* gene – was formally associated with global toxicity and HFS, as shown in Figure 4.9 and Table 4.11. Of note, the fine-maping SNP rs2741171 (r^2=0.73) – further downstream (27kb) of *TYMS* at chr18:700,687 – showed an even stronger association, reaching genome-wide significance for grade 01v2v34 HFS (OR=1.23, 95% CI=1.14-1.33, p=3.06e-08).

rs2612091 G-Allele	OR	95% CI	P-value
QUASAR2			
Global Binary: 012v34	1.59	(1.31-1.94)	4.01×10^{-6}
Global Quantitative: 01v2v34	1.19	(1.11-1.28)	1.68×10^{-6}
Diarrhoea Binary: 012v34	1.18	(0.88-1.60	0.27
Diarrhoea Binary: 01v234	1.04	(0.98-1.10)	0.24
HFS Binary: 012v34	1.59	(1.28-1.99)	3.46×10^{-5}
HFS Binary: 01v234	1.58	(1.31-1.91)	2.29×10^{-6}
HFS Quantitative: 01v2v34	1.21	(1.13-1.30)	2.59×10^{-7}
SCP			
HFS Binary: 012v34	1.64	(1.12-2.42)	0.012

Table 4.11: Associations with common 5-FU/capecitabine-related toxicity phenotypes for *TYMS/ENOSF1* rs2612091. No significant associations existed among the rare phenotypes or among other clinically relevant toxicity codings (eg, diarrhoea grade 01v234).

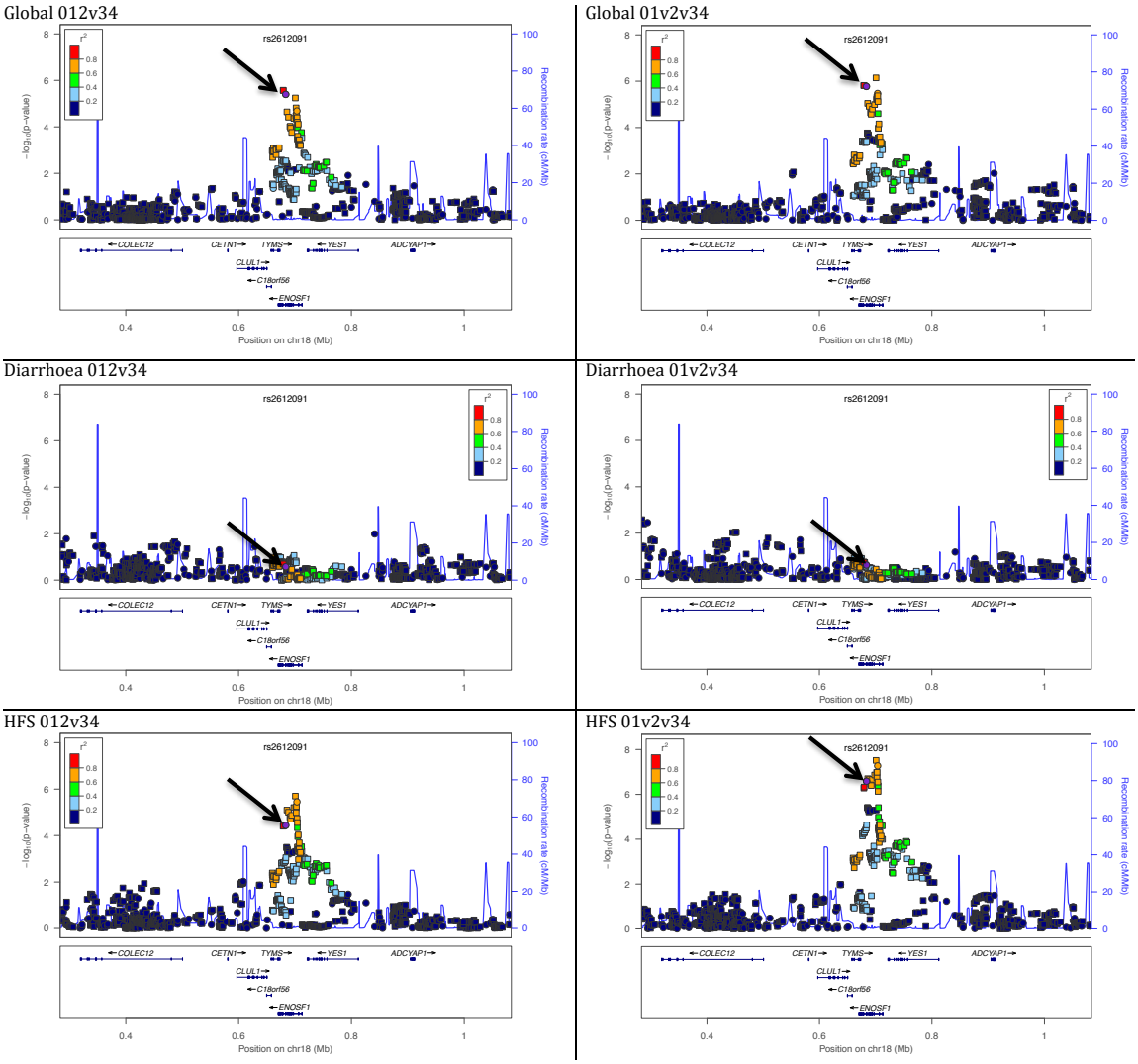


Figure 4.9: LocusZoom plots for associations with capecitabine/5-FU-related toxicity phenotypes for *TYMS/ENOSF1* rs2612091 and the surrounding region. The rs2612091 SNP, shown as a purple circle with an arrow pointed at it, is at chr1:683,607, 10kb downstream of the previously recognized putative 5-FU toxicity gene *TYMS* and intronic to the *TYMS* regulatory gene *ENOSF1*. Other circles represent associations among SNPs included on the directly genotyped 1,456 SNP/exome variant panel, while squares are fully imputed fine-mapping SNPs. Across the entire region, linkage disequilibrium is displayed by colour, with red SNPs having an r-squared of at least 0.8 with the lead SNP (rs2612091), orange having 0.6 to 0.8, green having 0.4 to 0.6, light blue having 0.2 to 0.4, and dark blue having little LD – all based on the March 2012 release of 1,000 Genomes. Recombination sites are marked by vertical purple lines and the location of *TYMS* and *ENOSF1* are shown below each plot. The SNP of highest significance in the global and HFS quantitative model is rs2741171.

4.5.6 rs2612091 Impacts *ENOSF1*, not *TYMS*, at the mRNA Expression Level

Both the rs2612091 and rs2741171 SNPs fall between recombination hotspots that flank the entirety of *TYMS* and *ENOSF1* (Figure 4.9). As shown in Figure 4.10, the *TYMS* and *ENOSF1* genes are partially antisense – they are coded on complementary strands of DNA and the 3'UTR of *ENOSF1* overlaps the 5'UTR and some exons of *TYMS* (53). Neither SNP, nor any of the highly correlated SNPs, has an obvious impact on either protein (ie, exonic or splice locus), with rs2612091 located 10kb downstream of *TYMS* between the 6th and 5th to last exons of *ENOSF1* and rs2741171 located 70kb downstream of *TYMS* between the 2nd and 3rd exons of *ENOSF1*. As such, while it was clear that neither variant directly altered the amino acid composition of either protein, it was not immediately obvious whether these SNPs affected levels of 5-FU/capecitabine-related toxicity via an impact on *TYMS* or an impact on *ENOSF1*, nor was it known whether the impact on toxicity resulted from changes at the transcriptional (eg, reduced mRNA formation from DNA) or translational (eg, reduced protein formation from mRNA) level, or both.

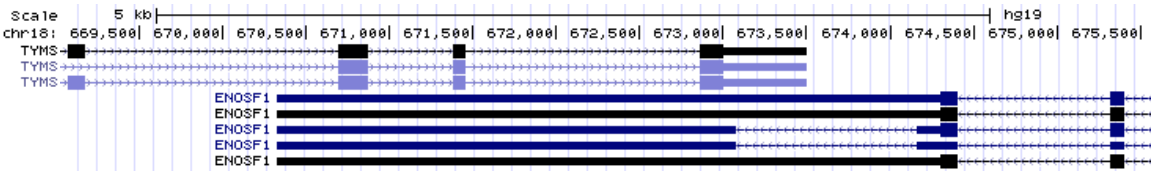


Figure 4.10: The *TYMS* and *ENOSF1* genes are antisense and partially overlapping, allowing for an RNA product of the *ENOSF1* untranslated region (UTR) that can bind to *TYMS* mRNA at the 3'UTR end. Both rs2612091 and rs2741171 are further right than shown on this image (downstream of *TYMS* and intronic to *ENOSF1*). Full thickness bands indicate exons, partial thickness bands indicate UTRs (transcribed into mRNA but not translated into protein) and horizontal lines indicate introns. Arrows indicate the direction of transcription, with left to right designated the positive DNA strand and right to left designated the negative (reverse complement). Image is a screenshot from the UCSC Bioinformatics Genome Browser, build 37.

To resolve the relationship between rs2612091 and *TYMS*/*ENOSF1*, we used public databases to assess the effect of rs2612091 – and, by correlation, rs2741171 – on

mRNA expression (ie, mRNA levels) of both *TYMS* and *ENOSF1*. No protein expression data were available. Using Caucasian matched twin data from Nica et al (146) as part of the Genevar database, we observed that the rs2612091 G-allele (high toxicity allele) significantly decreased *ENOSF1* mRNA expression in adipose tissue for both twin sets (p_{set1}=7.0e-04, p_{set2}=9.7e-06) and in lymphoblastoid cells for one twin set (p_{set1}=0.89, p_{set2}=0.0012), whereas rs2612091 genotype showed no association with *TYMS* mRNA expression (p>0.30 for each of the stated analyses). The result was similar using lymphoblastoid mRNA expression data from Stranger et al (188), in which the G-allele decreased mRNA expression of *ENOSF1* (p=1.9e-06) but not *TYMS* (p=0.82). Using data from Dimas et al, rs2612091 genotype had no effect on GenCord-derived fibroblast, lymphoblastoid or T-cell *TYMS* or *ENOSF1* mRNA levels (52). Finally, using eQTL data from colonic tissue in the TCGA dataset, the rs2612091 G-allele was again associated with decreased *ENOSF1* mRNA levels (OR=0.76, p=1.5e-07), but not with those of *TYMS* (OR=0.95, p=0.45). The implications of these findings are described in the Discussion section of this chapter.

4.5.7 rs2612091 Better Explains Toxicity Variation than do the *TYMS* 5'VNTR and/or 3'UTR Polymorphisms

While no protein expression data were available for the analysis of rs2612091, there is a clear association between two *TYMS* regulatory-region polymorphisms and 5-FU toxicity – the *TYMS* 5'VNTR and 3'UTR polymorphisms (see Chapter 3). While this might imply that rs2612091 functions via a TYMS-included mechanism, the effect of these polymorphisms on TYMS protein levels and activity remains unclear, as discussed in Introduction of this chapter. Given the absence of a direct chain from polymorphism to toxicity via *TYMS*, we tested whether the known *TYMS* 5'VNTR and 3'UTR polymorphisms might affect 5-FU/capecitabine toxicity as a result of correlation with rs2612091, which would partially discredit the putative role that

TYMS plays in affecting 5-FU toxicity by removing the gene's prior genetic basis for the phenotype.

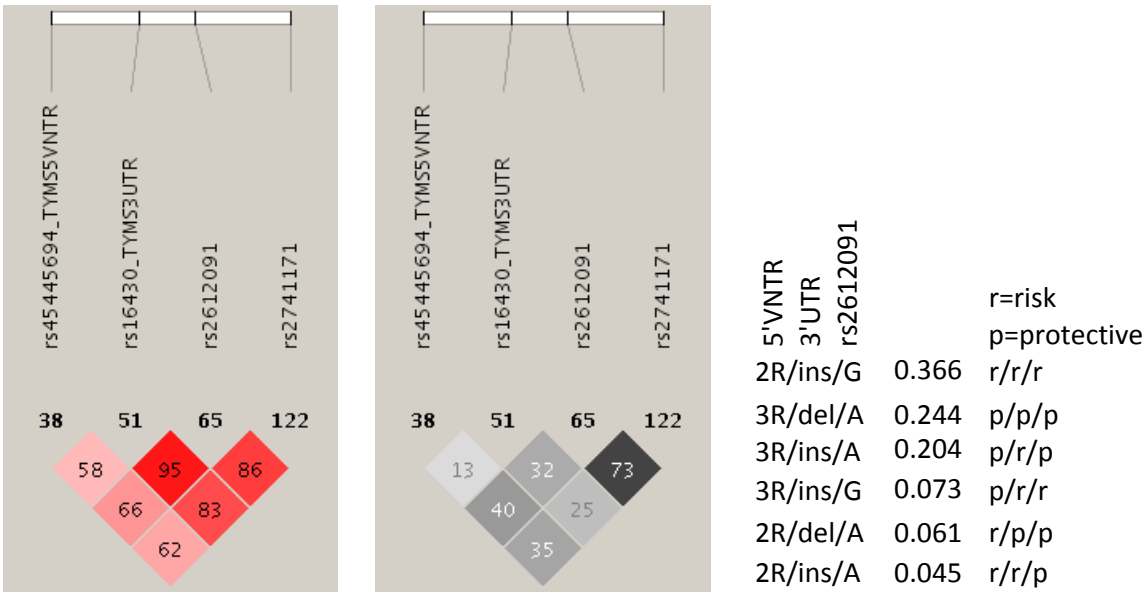


Figure 4.11: The linkage disequilibrium structure of the *TYMS/ENOSF1* variants was determined using QUASAR2 data in Haploview. The left-most image shows D' while the middle figure shows r^2 values. The right-most image shows the frequency of each 3-polymorphism haplotype, annotating which alleles are toxicity risk alleles.

As shown in Figure 4.11, the putative *TYMS* toxicity variants are not fully independent of each other or of rs2612091/rs2741171, but rather modest linkage disequilibrium is present. In order to test whether one, two or all three variants contribute to the increased risk of toxicity observed when each polymorphism is tested alone, the effects of these polymorphisms were tested in joint models, using handfoot syndrome as a quantitative variable (01v2v34) because this phenotype underlies the significant toxicity signal observed for all three variants (Section 4.5.5 and Chapter 3). First, the three polymorphisms were included in a covariate-adjusted logistic regression model. Only rs2612091 remained significantly associated with toxicity when adjusting for the genotype of the two *TYMS* UTR variants (rs2612091 alone, OR=1.2, p=4.2e-08; in 3-SNP model, rs2612091 OR=1.2, p=0.00049, 5'VNTR p=0.19, 3'UTR p=0.48). Further, as discussed in the 5-FU meta-analysis chapter (Section 3.9), the 5'VNTR

and 3'UTR variants both reduce in significance when included in a 2-SNP model, suggesting that both contribute to an individual's risk of toxicity. The lack of *TYMS* 5'VNTR/3'UTR significance suggested that rs2612091 captures the association signal resulting from all three variants; further, as rs2612091 alone minimises the model's AIC (Table 4.12), this SNP alone forms the best model for predicting toxicity risk affected by all three variants. Epistasis (interaction) among variants was also tested, and no evidence was found for rs2612091 and either the 5'VNTR (p=0.92) or 3'UTR (p=0.19), suggesting that the three variants do not act as a 3-variant tag for unidentified variants. Using the `-independent-effect` command in PLINK, the variants were then tested as a 3- and 2-polymorphism haplotype in which the effect of each test SNP is isolated by comparing the consistency of its effect across all combinations of the other two non-test SNPs. As shown in Table 4.13, in the 3-SNP model, while neither the 5'VNTR nor 3'UTR test allele has a consistent effect on HFS levels when varying the other two variants, the rs2612091 SNP's genotype is significantly associated with HFS levels (p=0.0021) when varying the genotype of the two *TYMS* UTR variants. The association becomes even more significant when comparing rs2612091 to the 5'VNTR and 3'UTR variants separately in a 2-SNP model.

SNP(s)	Global 012v34	HFS 01v2v34
5'VNTR/3'UTR/rs2612091	1075.4	2040.0
5'VNTR/rs2612091	1073.9	2038.4
3'UTR/rs2612091	1074.0	2039.3
rs2612091	**1072.5**	**2037.7**
5'VNTR/3'UTR	1082.1	2050.2
5'VNTR	1085.8	2049.6
3'UTR	1087.7	2060.7

Table 4.12: Akaike information criterion (AIC) values are shown for logistic regression models of toxicity that include between one and three *TYMS/ENOSF1* polymorphisms. For both the global binary (grade 012v34) and HFS quantitative (grade 01v2v34) models, the AIC is minimised when including only rs2612091 genotype information in the model. N=869 for all analyses. Clinical and genetic data are from QUASAR2.

Test SNP	5'VNTR 3'UTR rs2612091	OR for effect of each test SNP allele on haplotypes	OR for pooled effect of both test SNP alleles	p-value
5'VNTR	2R/ins/G	1 (ref)	1 (ref)	0.65
(3 SNP model)	3R/ins/G	1.04		
	2R/del/A	0.89	0.83	0.34
	3R/del/A	0.82		
	2R/ins/A	0.95	0.80	0.081
	3R/del/A	0.77		
	overall			0.17
3'UTR	2R/ins/G	1 (ref)	1 (ref)	n/a
(3 SNP model)	3R/ins/G	1.04	1.04	n/a
	2R/del/A	0.89	0.92	0.67
	2R/ins/A	0.95		
	3R/del/A	0.82	0.80	0.33
	3R/del/A	0.77		
	overall			0.61
rs2612091	2R/ins/G	1 (ref)	1 (ref)	0.66
(3 SNP model)	2R/ins/A	0.95		
	3R/ins/G	1.04	0.84	0.00068
	3R/ins/A	0.77		
	2R/del/A	0.89	0.88	n/a
	3R/del/A	0.82	0.82	n/a
	overall			0.0021
rs2612091	2R/G	1 (ref)	1 (ref)	0.18
(2 SNP model)	2R/A	0.92		
	3R/G	1.04	0.84	0.00051
	3R/A	0.79		
	overall			0.00053
rs2612091	ins/G	1 (ref)	1 (ref)	
(2 SNP model)	ins/A	0.80		n/a
	del/A	0.83	0.90	n/a
	overall			1.47E-06

Table 4.13: The independence of rs2612091 from the two *TYMS* UTR putative toxicity variants was assessed by isolating the effect of each allele irrespective of genotype for the non-test SNPs. Two types of p-values are produced – (i) one for the effect of each test-SNP test allele on toxicity, given one particular combination of the non-test SNP alleles, and (ii) one for the effect of the test-SNP's test allele on toxicity, given all genotype combinations of the non-test SNP ("overall"). Note that some rare haplotypes are not shown.

Considering the independent effect results of this section together with the mRNA expression levels of the previous section and the literature review in Introduction, the role that polymorphisms in or near to the *TYMS* gene play in 5-FU/capecitabine-related toxicity is discredited, giving us less prior basis for assigning the toxicity impact of rs2612091 to a protein effect in *TYMS* over *ENOSF1*. Rather, it appears that rs2612091 alone can explain the relationship between the *TYMS* UTR polymorphisms and 5-FU/capecitabine-related toxicity, suggesting that the 5'VNTR and 3'UTR polymorphisms may act as tags for variation that directly affects products of the *ENOSF1* gene, particularly mRNA levels.

4.5.8 Set-Based Tests of Capecitabine/5-FU Pathway Genes

As discussed in Chapter 3, power calculations for individual SNP associations showed that the QUASAR2 sample size and event rate allowed approximately 90% power to detect an association with global toxicity for either (i) an odds ratio of >1.5 for variants with a MAF of 10% or (ii) an odds ratio of >3.0 for variants with a MAF of 1% (see further calculations in Appendix). As such, our analysis was potentially underpowered for the identification of individual SNPs with rare MAF and modest effect size. To help assess whether entire genes, or even the collection of genes, in the 5-FU/capecitabine pathway contained multiple variants for which we had insufficient statistical power for their individual identification, set-tests were performed. Specifically, the `-set-test` command in PLINK was used to assess whether the distribution of allelic association p-values observed among all SNP/exome variants in a given gene (and then the pathway as a whole) was more significant than what would be expected by chance (ie, randomly assigned phenotypes). For the analysis, the previously-reported toxicity variants and the three variants identified by this study were excluded, as were SNPs with r^2>0.1 LD with these SNPs. As shown in Table 4.14, although some individual genes showed modest significance (eg, *TYMP* for both

HFS and diarrhoea), none was considered significant upon applying a false discovery rate significance threshold.

SET	No. SNPs	Global 012v34	Global 01v2v34	HFS 012v34	HFS 01v2v34	Diarrhoea 012v34	Diarrhoea 01v2v34
ABCB1	77	0.29	1	1	1	1	1
ABCC3	64	1	1	1	1	0.55	0.49
ABCC4	221	0.86	0.87	1	1	0.91	0.75
ABCC5	100	0.33	1	1	1	0.13	0.074
ABCG2	57	0.093	1	0.37	0.60	1	1
CDA	25	1	0.24	0.38	0.20	1	0.40
CES1	24	1	0.10	0.19	0.11	1	0.16
CES2	59	0.18	0.28	0.0092	0.046	1	1
DPYD	189	0.31	0.040	0.52	0.52	0.53	0.20
DPYS	69	1	0.32	1	0.58	1	0.70
ENOSF1	22	0.11	0.27	1	0.25	1	1
MTHFR	37	1	1	1	1	1	0.50
PPAT	29	1	1	1	1	1	0.10
RRM1	29	1	1	0.17	1	1	1
RRM2	19	1	1	1	1	1	1
SLC22A7	26	1	1	1	1	0.014	0.22
SLC29A1	26	1	1	1	1	1	1
TK1	35	1	1	1	0.43	1	1
TYMP	92	1	0.24	0.035	0.025	0.055	0.029
TYMS	23	0.12	0.28	1	0.25	1	1
UCK1	43	0.13	0.16	0.25	0.36	0.013	0.0038
UCK2	21	1	1	1	1	0.063	0.31
UMPS	34	0.13	0.085	0.15	0.035	1	0.32
UPB1	30	1	1	1	1	1	1
UPP1	16	1	1	1	1	0.089	1
UPP2	42	1	0.061	0.078	0.25	1	1
As One Set*	**1393**	**0.72**	**0.36**	**0.67**	**0.55**	**0.10**	**0.12**

Table 4.14: SNP and exon array variants were assessed by permuting the phenotype data 10,000 times and comparing the observed distribution of p-values to that obtained by chance. A FDR-adjusted significance threshold of p=.0019 was used to determine statistical significance.

4.5.9 *DPYD* A551T is a Novel Rare 5-FU/Capecitabine Toxicity Variant

In addition to the SNP and exome array variants described in the previous subsections, this chapter contains results for the 100 high toxicity ("HT") and 100 low

toxicity ("LT") QUASAR2 patients compared for an allele frequency difference at exonic and regulatory regions of *DPYD* and *TYMS* as part of the Roche/454 sequencing project. As discussed in Section 4.4, within the target regions for *DPYD*, *TYMS* and *UPP2*, 39 variants were called and passed quality control filtering. Figure 4.12 shows that these 39 variants were spread across the three genes and that all but 18 were either analysed in the full QUASAR2 patient set as part of the SNP or exome array analysis of this chapter or are a known putative 5-FU toxicity variant analysed as part of the meta-analysis in the previous chapter.

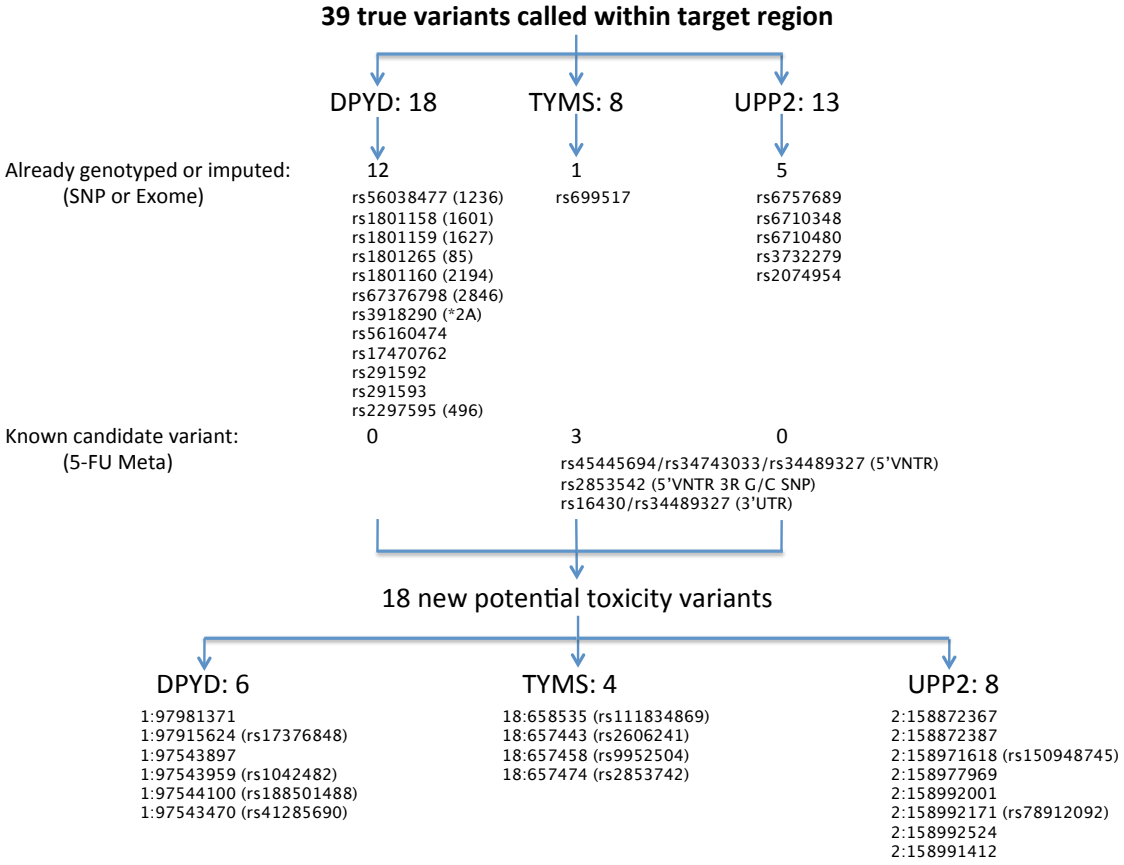


Figure 4.12: Among the true variants identified in Figure 4.6, many were already analysed for an association with 5-FU/capecitabine toxicity as part of the meta-analysis and/or SNP and exome array work described in this chapter. After removing previously tested variants, 18 potential new toxicity variants remained for analysis and follow-up.

The list of 18 remaining variants – a mixture of dbSNPs and new SNPs – were then assessed by comparing the frequency of the minor allele in the high toxicity and low

toxicity patient pool and by functional annotation with ANNOVAR, as shown in Table 4.15. Among those polymorphisms with an allele frequency difference between the 100 LT and 100 HT patients, follow-up analysis was an option via genotyping individuals in the entire QUASAR2 cohort. Selection of 5-FU pathway genes for sequencing was based on association testing prior to imputation. Given time and reagent expense, follow-up was deprioritised for the *UPP2* variants when the original tagging SNP association used to select *UPP2* lost candidate gene level significance upon fill-in imputation of the QUASAR2 dataset. The *TYMS* and *DPYD* variants maintained significance (Section 4.5). Among *TYMS* variants identified, all were in the CpG island and while these are interesting candidate loci for further analysis, they again were deprioritised based on time and reagent costs. Among the *DPYD* variants, one missense mutation (A551T) showed an approximately 9-fold higher allele frequency in the HT pool than in the LT pool. This variant was predicted to be strongly damaging by SIFT, Polyphen, Phylop, and MutationTaster. Sanger sequencing of this variant in the full HT pool identified a single patient who held the rare allele, and this patient experienced grade 4 neutropaenia and thrombocytopaenia – considered the most clinically severe adverse events apart from death. As shown in Table 3.10 of the meta-analysis chapter, this variant has previously been identified as a causal DPYD Deficiency Syndrome variant (205), though the variant has not been validated by *in vitro* studies of DPYD activity, such as by Offer et al (149). This variant, along with over 15 other causal DPYD Deficiency Syndrome variants, are strong candidates for follow-up genotyping among large patient sets, though functional work may prove a better means of analysis, given their low frequency in the population.

Gene	b37 ChrCoordinate	rsID	Func	Exon Func	High Tox Var Allele Reads	High Tox WT Allele Reads	Low Tox Var Allele Reads	Low Tox WT Allele Reads	High/Low Tox Read Count OR	FisherP	PearsonP	MAF (Observed)	Follow-up Genotyping
DPYD	1:97981371		exonic	G1651A; A551T	90	15258	8	12626	9.4	1.4E-15	1.7E-13	0.35%	Deprioritised: rare
DPYD	1:97915624	rs17376848	exonic	T1896C; F632F	979	22988	466	15878	1.5	3.7E-11	6.1E-11	3.58%	Deprioritised: synonymous
DPYD	1:97543897		3'UTR		111	12896	0	5162	Inf	1.3E-16	2.8E-11	0.61%	Deprioritised: rare
DPYD	1:97543959	rs1042482	3'UTR		774	11550	374	4433	0.8	5.2E-04	4.2E-04	6.70%	Deprioritised: proxy not assoc.
DPYD	1:97544100	rs188501488	3'UTR		300	19761	18	8203	6.9	5.1E-26	2.4E-20	1.12%	Deprioritised: UTR
DPYD	1:97543470	rs41285690	3'UTR		50	6017	165	5482	0.3	9.8E-18	2.9E-17	1.84%	Deprioritised: UTR
TYMS	18:658535	rs111834869	CpG indel		2159	17312	847	7280	1.1	1.1E-01	1.1E-01	10.89%	Deprioritised: no MAF diff
TYMS	18:657443	rs2606241	CpG		1890	8223	2363	3157	0.3	1.3E-223	4.2E-230	27.21%	Deprioritised: CpG
TYMS	18:657458	rs9952504	CpG		390	9547	290	4837	0.7	1.8E-06	1.2E-06	4.52%	Deprioritised: CpG
TYMS	18:657474	rs2853742	CpG		3341	6028	1525	2703	1.0	6.6E-01	6.5E-01	35.79%	Deprioritised: no MAF diff
UPP2	2:158872367		exonic	C116G; P39R	206	9330	134	8919	1.5	5.5E-04	5.4E-04	1.83%	Deprioritised: no SNP assoc.
UPP2	2:158872387		exonic	G136T; A46S	1	9520	68	9007	0.0	2.8E-20	1.2E-16	0.37%	Deprioritised: no SNP assoc.
UPP2	2:158971618	rs150948745	exonic	C186G; V62V	233	11702	26	5811	4.4	6.9E-18	3.5E-15	1.46%	Deprioritised: no SNP assoc.
UPP2	2:158977969		exonic	T503G; F168C	19	3462	0	12761	Inf	1.9E-13	6.8E-17	0.12%	Deprioritised: no SNP assoc.
UPP2	2:158992001		3'UTR		0	12503	162	9209	0.0	1.0E-60	2.8E-49	0.74%	Deprioritised: no SNP assoc.
UPP2	2:158992171	rs78912092	3'UTR		111	12610	0	9709	Inf	5.0E-28	2.8E-20	0.50%	Deprioritised: no SNP assoc.
UPP2	2:158992524		3'UTR		44	12138	41	7126	0.6	4.2E-02	3.2E-02	0.44%	Deprioritised: no SNP assoc.
UPP2	2:158991412		3'UTR		4	9376	55	7793	0.1	6.7E-15	1.8E-13	0.34%	Deprioritised: no SNP assoc.

Table 4.15: The 18 novel putative toxicity variants in *DPYD*, *TYMS* and *UPP2* were assessed by comparing allele frequency of patients in the high toxicity and low toxicity pool and by functional annotation. Among the one missense *DPYD* variant and the three missense *UPP2* variants, *DPYD* A551 showed strong evidence of functional impact by SIFT, Polyphen, Phylop, and MutationTaster, while only *UPP2* F168C showed modest support from these annotation tools.

4.5.10 ROC Analysis of New and Known 5-FU/Capecitabine Toxicity Variants in QUASAR2

In order to test whether the newly identified variants improved upon the known variants for the prediction of capecitabine-induced toxicity, *DPYD* 2846T>A, *DPYD* *2A, *DPYD* rs12132152, *DPYD* rs7548189, *DPYD* A551T, and *TYMS/ENOSF1* rs2612091 were incorporated into a ROC model in QUASAR2. The same method as described in 3.16 was used, applying the beta coefficients from the Caronia et al study where available and using 938 patients. Again, the (three) rare *DPYD* alleles were considered functionally equivalent and combined in a group test of any rare functional variant vs no rare allele (OR=7.6, p=4.5e-04). Although this test will be somewhat biased toward a higher score based on testing in its discovery set, the area under curve (AUC) improved from approximately 0.60 when using just the meta-analysis set of variants to 0.66 (95% CI 0.63-0.70) when incorporating the new panel, as shown in Figure 4.13. Specifically, at the cut-off for which the maximum proportion of patients was correctly classified (69%), sensitivity was 27% (95% CI 23-33%), specificity was 91% (95% CI 88-93%), positive predictive value was 60% (PPV: 95% CI 52-68%), and negative predictive value was 71% (NPV: 95% CI 68-74%).

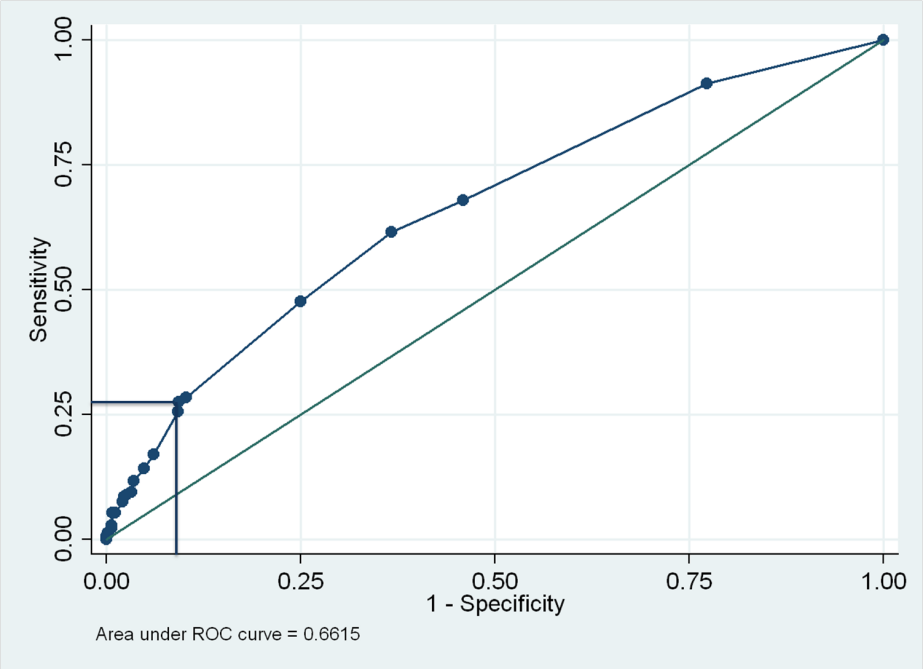


Figure 4.13: ROC (receiver operator characteristic) curve analysis was performed on 938 QUASAR2 patients using the variants identified in this chapter and in the meta-analysis of the previous chapter.

4.6 Discussion

This chapter presented technical and genetic results for the study of capecitabine-related toxicity in and around genes that encode enzymes involved in the drug's activation, action and catabolism. Three novel associations with capecitabine/5-FU toxicity were identified in *DPYD* (rs12132152, rs7548189, A551T) and the previously identified *TYMS* 5'VNTR and 3'UTR toxicity polymorphisms were refined to a tagging SNP (rs2612091) downstream of *TYMS* and intronic to the adjacent *ENOSF1*.

The most surprising result was that the only two *TYMS* variants previously shown to impact toxicity risk – the 5'VNTR and 3'UTR polymorphisms – appear to not play a functional role in affecting levels of 5-FU toxicity, but rather they tag the effect of genetic variation elsewhere. This was identified through the discovery that rs2612091, a tag SNP 10kb downstream of *TYMS* in an intron of the adjacent *ENOSF1* gene, not only captures the effect on toxicity of the 5'VNTR and 3'UTR polymorpshisms,

but also shows additional effect not captured by these *TYMS* polymorphisms. The relationship with toxicity was even stronger for the rs2741171 SNP, 17kb further downstream of *TYMS*.

To resolve the relationship between rs2612091/rs2741171 and the *TYMS* and *ENOSF1* genes, we used public databases to investigate whether rs2612091 affects mRNA levels of these two genes. Given that the *ENOSF1* alpha transcript (36) and beta protein (54) have been shown to decrease *TYMS* mRNA and protein expression, respectively, and that higher levels of *TYMS* mRNA (160) and protein (112) have been associated with decreased 5-FU toxicity, we expected that the rs2612091 G allele (ie, toxicity risk allele) would be associated with increased *ENOSF1* mRNA levels, such that the G allele would reduce down-regulation (ie, degradation) of *TYMS* mRNA. The results from multiple tissue types from multiple databases showed otherwise, with the rs2612091 G allele decreasing *ENOSF1* mRNA levels while increasing toxicity, suggesting that the SNP does not act through antisense-mediated transcript effects on *TYMS*. Given the results presented in this chapter and the controversy regarding the functional impact of the *TYMS* 5'VNTR and 3'UTR polymorphisms (69), it appears that the 5'VNTR and 3'UTR variants impact 5-FU toxicity through correlation with rs2612091 and the polymorphisms this SNP tags. As a start, it would be useful to test whether the relationships at the mRNA level remained at the TYMS and ENSOF1 protein level, or whether lower *ENOSF1* mRNA levels somehow resulted in lower TYMS protein expression. Previous authors have identified ENOSF1 proteins, specifically in a study that noted increased expression within 5-FU resistant cells (55).

We were also surprised to find that both *DPYD* rs12132152 failed to replicate its association with HFS in the SCP study and rs7548189 failed to replicate its association with diarrhoea, though the latter did replicate for its association with HFS. The failure of rs12132152 to replicate is particularly striking, as the SNP achieved near

genome-wide significance in the QUASAR2 analysis. Possibly, this suggests that the functional variants tagged by this SNP differ between the northern and southern European populations; however, the differences in patient recruitment between the SCP and QUASAR2 study might also account for the differences. A suitable UK-based study would provide a useful dataset for replication of this association.

The results of the Roche/454 sequencing project were promising, with one new toxicity variant identified and a short list of additional variants to consider for follow-up genotyping. The identification of this variant highlights the success of the experiment from a technical standpoint, as we were able to detect an allele difference in the high and low toxicity pools for a variant held by one heterozygous individual. For simplicity, future experiments of many amplicons and many patients could use bar codes that would tag the DNA of individual patients.

While the newly identified polymorphisms, used alongside known toxicity variants from Chapter 3, improve upon our ability to predict capecitabine-related toxicity using genetic biomarkers, such a panel has not yet achieved broad clinical utility. Further, the set-test results from this chapter suggest that the lack of additional SNP associations was not the result of a partially powered study for the identification of variants with somewhat lower frequency or smaller effect size. As such, Chapter 5 analyses the SNP and exome array variants on a genome-wide basis.

Chapter 5

Capecitabine Toxicity GWAS of QUASAR2 and SCP Studies

5.1 Introduction

5.1.1 Background

The candidate gene analysis presented in Chapter 4 identified four novel SNP associations with 5-FU/capecitabine-related toxicity within the drug's activation, target and degradation pathway. Despite dense common and rare SNP coverage, this analysis did not identify enough biomarkers so as to create a clinically useful predictive kit, suggesting that variants beyond this pathway may be involved. While a drug pathway-based analysis allows for a focused investigation of SNPs with higher likelihood of affecting toxicity, expanding the analysis genome-wide allows for the discovery of additional variants that might provide novel insights into the mechanisms that result in toxicity to 5-FU/capecitabine. Such variants would also improve our ability to predict toxicity to the drug.

For the most part, GWAS has neither provided a clinically useful set of markers for the prediction of toxicity nor has it greatly deepened our understanding of a drug's

pharmacokinetics or pharmacodynamics. Seven GWAS have been performed on toxicity to cancer chemotherapy – for children with acute lymphoblastic leukemia given methotrexate (198), for musculoskeletal adverse events from aromatase inhibitors (96), for extreme high/low toxicity patients given epirubicin (185), for gemcitabine-induced events (113), for irinotecan-related diarrhoea and neutropaenia in non-small-cell lung cancer patients (81), for FOLFOX-related toxicity in CRC patients (63), and for sensory peripheral neuropathy resulting from paclitaxel (14). Only one has reported a genome-wide (p<5e-08) significant result (198), though the association (p=1.7e-10) was between genotype and drug clearance and the associations with toxicity were less significant (p=0.004 and above). The association, however, did identify a drug transporter that had never previously been studied as part of the drug's pathway.

Most of these studies have had low sample size (discovery: range=79-878, median=270, mean=405; replication: range=0-791, median=146, mean=222) and given the success of GWAS in identifying variants associated with disease predisposition (Section 1.4.3) – and, to some degree, its success in identifying variants associated with survival (97) and safe/efficacious dosing (41) – it remains probable that the method is appropriate for the identification of novel variants and unknown actors in the 5-FU/capecitabine toxicity story, particularly when using a large sample size and reliable toxicity data.

5.1.2 Aims and Overview

Using over 1,000 patients from the QUASAR2 trial and genome-wide meta-analysis with data from the SCP study (Section 4.1.3), this chapter performs a GWAS of common and rare SNPs genotyped using the methods described in Chapter 4. Additionally, the SNP data from 940 QUASAR2 patients are used to estimate the heritable component of inter-individual toxicity variation and determine whether associated

variants impact risk of particular toxicity phenotypes independently from risk for other phenotypes. While further replication is needed to validate the strongest associations found in these analyses, this chapter provides a short list of variants to prioritise for follow-up study.

5.2 Heritability of Capecitabine-Related Toxicity

An estimate of the heritability of CRC toxicity found within common alleles was performed using LDAK software (183) to calculate a weighted kinship matrix of SNP independence and restricted maximum likelihood (REML) analysis using GCTA software (218). Patient age, patient gender and treatment arm were incorporated into the analysis as covariates, as were the first 10 principal components from PCA, as population structure differences among cases and controls can cause an inaccurate prediction (28). As shown in Figure 5.1, the 940 QUASAR2 patients used throughout this chapter (those on the SNP arrays and imputed) cluster very tightly when plotting the first and second eigenvector in an X/Y space (no PCs were significant at $p<0.05$ for global toxicity in QUASAR2), suggesting low risk of population structure differences resulting in spurious heritability estimates or (later) association tests. Although a handful of QUASAR2 samples cluster less tightly than the rest of the cases/controls, these patients were all controls and were included in the analysis because they generally clustered as CEU and were not outliers when analysed as part of a broader set of UK patients in the genotype data's original quality control measures.

As shown in Table 5.1, heritability estimates of 5-FU/capecitabine-related toxicity were performed for global toxicity, diarrhoea and HFS. In none of these analyses was the heritable component estimated to be significantly greater than 0%. However, none of the analyses appeared to be sufficiently statistically powered, as the confidence intervals on the 0% heritability estimates ranged as high as 75% heritable. The

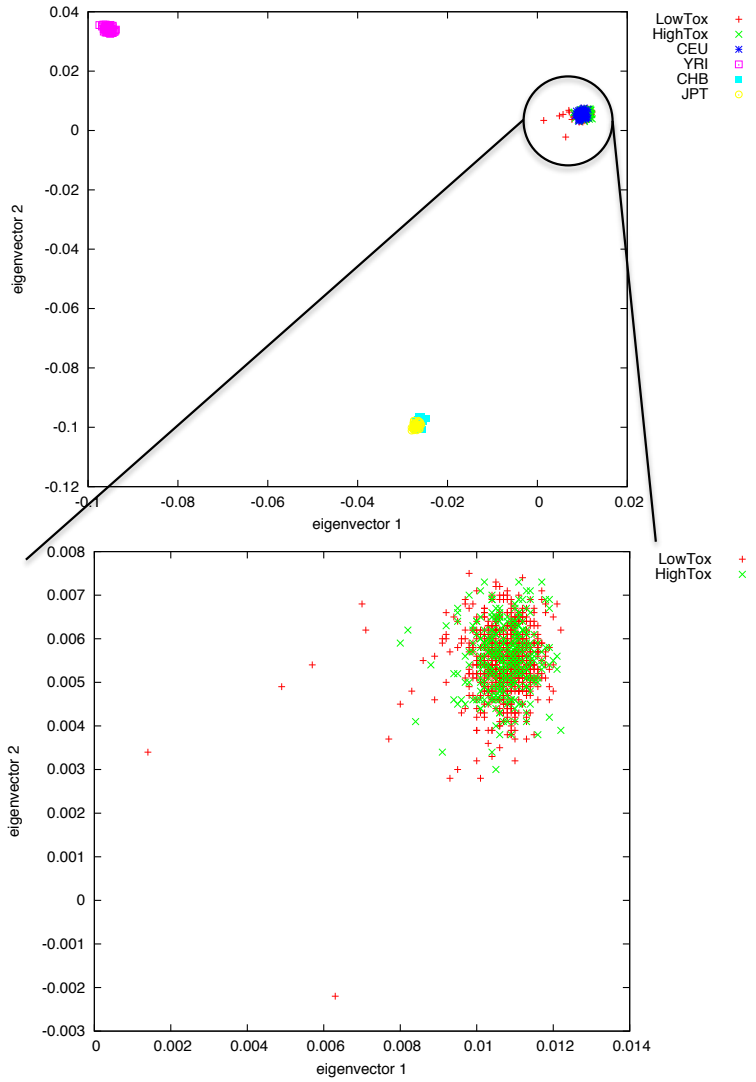


Figure 5.1: The first and second eigenvector of the principal components analysis of the 940 QUASAR2 patients included in the various analyses of this chapter are shown in an X/Y plot. QUASAR2 global toxicity grade 0/1/2 controls are shown in red +s, while cases are shown in green Xs. The top plot shows how the cases and controls cluster among 268 HapMap population controls, while the bottom plot zooms in on the sample cluster and includes only QUASAR2 cases/controls.

results were similar when applying quantitative and binary codings to the toxicity data. This result was partially expected, as the sample size was 940 and a sample size of 2,000 is the suggested minimum. The result, however, does not mean that individual biomarkers of 5-FU toxicity do not exist, as is evidenced by the fact that the significant markers from Chapter 4 were included in the heritability estimates.

		Variance	SE	
Global Toxicity (Quantitative)	**Covariates**	Age, gender, arm		
	Vg	0.00%	0.324	
	Ve	82.17%	0.324	
	Vp	82.17%	0.038	
	Vg/Vp	**0.00%**	0.394	
	p-value			0.5
Diarrhoea (Quantitative)	**Covariates**	Age, gender, arm		
	Vg	0.00%	0.371	
	Ve	92.70%	0.373	
	Vp	92.70%	0.043	
	Vg/Vp	**0.00%**	0.400	
	p-value			0.5
HFS (Quantitative)	**Covariates**	Age, gender, arm		
	Vg	0.44%	0.401	
	Ve	101.76%	0.404	
	Vp	102.21%	0.048	
	Vg/Vp	**0.43%**	0.393	
	p-value			0.5

Table 5.1: Quantitative outcome variables for 5-FU/capecitabine-related toxicity were tested for a heritable component using the LDAK/GCTA method. The Vg/Vp value is an estimate of the proportion of phenotypic variance that is attributable to genetic variability. Estimates have been transformed to percentages, whereas standard errors (SE) of these estimates have not. Thus, assuming a normal distribution, the confidence of the estimate would be the raw estimate +/- 1.96*SE, which produces a 95% confidence interval from 0% to approximately 75% heritability for all estimates. Vg=genetic variance estimate, Ve=residual variance (environmental), Vp=phenotypic variance, Vg/Vp=ratio of genetic variance to phenotypic variance. P-values are from a log-likelihood ratio on 1 degree of freedom.

5.3 GWAS of Capecitabine-Related Toxicity in the QUASAR2 Study

This chapter investigates genome-wide the SNP array, imputation and exome array datasets described in Section 4.2 and Section 4.3 of the Candidate Gene Analysis chapter. As in Chapter 4, 1,046 QUASAR2 patients were included – 12 in only the SNP array analysis, 102 in only the exome array analysis and the remainder in both – for both analyses of global toxicity and for analyses of each of the six specific

5-FU-related toxicity phenotypes comprising the global measure. Common phenotypes (global, diarrhoea and HFS) were tested under both binary (grade 012v34) and quantitative (grade 01v2v34) models, while rare phenotypes were tested only under a binary (grade 0v1234) model. Instead of extracting only 1,456 capecitabine-pathway genetic variants for analysis, this chapter analyses the full genetic dataset, including 1,044,711 directly genotyped variants from the Hap300/370, Hap610 and/or Omni2.5 SNP arrays, which were imputed to 4,777,723 SNPs for all of the 940 QUASAR2 patients genotyped on any of these three platforms. Of these, 4,706,087 SNPs had a SNPTEST info score >0.95 for imputation accuracy and a MAF>0.01 and were analysed under an allelic model in SNPTEST (methods in Section 2.4). In addition, 247,703 rarer and mostly coding variants on the Illumina Human Exome Array platform were analysed under both an allelic and dominant model using PLINK, of which 232,580 were non-overlapping with the imputed SNP array content and were included in the analysis.

Top associations from the QUASAR2-only SNP and exome array variant analysis are shown in Table 5.2, which includes all associations at p<1e-06 and associations for nonsynonymous variants at p<5e-06. No SNPs achieved this threshold for either the binary or quantitative global toxicity measure. Manhattan and QQ plots of the genome-wide analysis are shown in Section 5.4 as some combine the outputs from the QUASAR2-only analysis and QUASAR2/SCP meta-analysis.

In addition to the *DPYD* and *TYMS/ENOSF1* variants described in Chapter 4, one variant reached genome-wide significance among all toxicities/models tested: rs2093152, an intergenic SNP 30kb from *TSHZ2* (*Homo sapiens* teashirt zinc finger homeobox 2) and 750kb from *ZFP64* (*Homo sapiens* ZFP64 zinc finger protein). Visualising the ENCODE regulation tracks available from UCSC, this variant sits in a histone H3 lysine 4 tri-methylation site, suggesting it might be involved in promoter activity (87). Given the broad roles played by zinc fingers, it is hard to pinpoint the

							Primary Toxicity Association				Global 012v34 Toxicity								
Toxicity	Coding	Model	Variant	Chr	Coord	Ref	Test	OR	LCI	RCI	P-value	OR	LCI	RCI	P-value	Info	MAF	N	Imp	Function	Gene
Diarrhoea	012v34	DOM	exm21596/rs221057	1	16534255	C	G	0.31	0.19	0.51	4.37E-06	0.73	0.58	0.92	6.81E-03	NA	0.25	966	0	NonSynon	ARHGEF19
Diarrhoea	012v34	ADD	rs4309833	4	10666766	A	G	0.45	0.33	0.60	1.33E-07	0.78	0.64	0.95	1.19E-02	0.99	0.42	940	940	intronic	CLNK
Diarrhoea	012v34	ADD	rs75153039	4	149609427	A	C	4.69	2.58	8.52	4.05E-07	1.18	0.80	1.74	4.06E-01	0.99	0.07	940	940	intergenic	NR3C2 (dist=245755), DCLK2 (dist=1389999)
Diarrhoea	012v34	ADD	rs78285023	7	121488024	T	C	23.35	6.83	79.85	5.22E-07	2.35	1.08	5.12	3.14E-02	0.97	0.01	940	940	intergenic	FAM3C (dist=451602), PTPRZ1 (dist=25135)
Diarrhoea	01v2v34	ADD	rs78285023	7	121488024	T	C	1.94	1.52	2.47	8.34E-08	2.35	1.08	5.12	3.14E-02	0.97	0.01	940	940	intergenic	FAM3C (dist=451602), PTPRZ1 (dist=25135)
Diarrhoea	012v34	ADD	rs7861903	9	128282914	G	A	7.33	3.31	16.21	8.86E-07	1.37	0.81	2.32	2.35E-01	0.98	0.04	940	484	intronic	MAPKAP1
Diarrhoea	01v2v34	DOM	exm1048238/rs10773019	12	124104551	A	G	1.39	1.21	1.61	4.41E-06	1.23	0.80	1.88	3.45E-01	NA	0.05	966	0	NonSynon	DDX55
Diarrhoea	012v34	ADD	rs2093152	20	51559324	G	C	10.55	4.66	23.90	1.60E-08	2.48	1.44	4.28	1.06E-03	0.96	0.03	940	940	intergenic	ZFP64 (dist=750800), TSHZ2 (dist=29622)
Diarrhoea	01v2v34	ADD	rs2093152	20	51559324	G	C	1.64	1.38	1.95	1.61E-08	2.48	1.44	4.28	1.06E-03	0.96	0.03	940	940	intergenic	ZFP64 (dist=750800), TSHZ2 (dist=29622)
Mucositis	0v1234	ADD	rs73029669	11	127424650	C	T	5.11	2.73	9.54	3.13E-07	0.68	0.37	1.24	2.03E-01	0.98	0.03	940	940	intergenic	KIRREL3-AS3 (dist=548697), ETS1 (dist=904006)
Mucositis	0v1234	ADD	rs12950574	17	73572478	T	C	1.72	1.39	2.14	7.55E-07	1.21	0.98	1.49	7.49E-02	0.99	0.34	940	940	intergenic	LLGL2 (dist=1188), MYO15B (dist=11661)
Neutropaenia	0v1234	ADD	SNP6-36830174	6	36722196	A	G	2.88	1.92	4.32	3.52E-07	1.07	0.82	1.40	6.27E-01	1.00	0.15	940	940	intronic	CPNE5
Neutropaenia	0v1234	ADD	SNP8-39956402	8	39837245	G	A	0.41	0.29	0.59	9.28E-07	0.87	0.69	1.10	2.57E-01	0.95	0.23	940	940	intronic	IDO2
Platelets	0v1234	ADD	rs315721	5	169715038	A	G	2.68	1.83	3.91	3.66E-07	0.93	0.75	1.15	4.89E-01	1.00	0.30	940	940	intronic	LCP2
Platelets	0v1234	ADD	rs3104945	8	96479283	A	G	13.40	4.82	37.29	6.62E-07	1.04	0.58	1.87	8.88E-01	0.96	0.03	940	940	ncRNA intronic	LOC100616530
Platelets	0v1234	ADD	SNP9-111682561	9	112642740	G	A	11.10	4.28	28.76	7.35E-07	0.64	0.37	1.10	1.04E-01	1.00	0.03	940	940	intronic	PALM2, PALM2-AKAP2
Platelets	0v1234	ADD	exm1047255/rs58843120	12	123444507	G	T	7.73	3.27	18.28	3.23E-06	2.35	1.10	5.03	2.72E-02	NA	0.01	965	940	NonSynon	ABCB9
Platelets	0v1234	ADD	rs34521141	20	50434898	A	C	3.76	2.23	6.35	7.16E-07	1.18	0.88	1.60	2.66E-01	0.98	0.12	940	940	intergenic	SALL4 (dist=15850), ZFP64 (dist=265652)

Table 5.2: Results from the QUASAR2-only analysis of common and rare SNP/exome array and imputed variants are shown. 16 variants had associations with one of the tested 5-FU/capecitabine-related phenotypes/codings at p<1e-06 or p<5e-06 for nonsynonymous variants, while none were significant for global toxicity. Global binary grade 012v34 toxicity results for these 16 variants are also shown. Note that exm21596/rs221057 is in perfect LD with another nonsynonymous variant in *ARHGEF19*, exm21605/rs221058 at chr1:16534646. The most significant SNP, rs2093152, contains only one directly genotyped proxy, rs4356719, but this SNP was only genotyped for 92 patients in QUASAR2. In addition, rs2612091/rs2741171, near *TYMS/ENOSF1*, and rs12132152/rs76387818, near *DPYD*, were significant at p<1e-06 for specific toxicity phenotypes, as described in Chapter 4. N=number of analysed patients; Imp=number of patients with imputed genotypes; NonSynon=Nonsynonymous single nucleotide variant. Variants with "exm" identifiers listed were analysed using genotype data from the exome arrays, whereas other data were from the SNP arrays (including imputation).

exact mechanism by which promoter activity relating to either of the adjacent genes would affect toxicity levels.

A few of the top associations stand out based on the function of their nearest-by gene(s). The rs75153039 variant was associated with diarrhoea and exists 246kb from *NR3C2* (nuclear receptor subfamily 3, group C, member), which mediates salt/water balance. The rs34521141 SNP was associated with thrombocytopaenia/platelets and exists 16kb from *SALL4* (*Homo sapiens* sal-like 4), another zinc-finger transcription factor, mutations in which are causative of IVIC syndrome (OMIM 147750), which is related to Duane-Radial Ray Syndrome but also causes mild thrombocytopaenia (9). The exm1047255/rs58843120 variant, also associated with thrombocytopaenia, is a missense single nucleotide variant within *ABCB9* (*Homo sapiens* ATP-binding cassette, sub-family B (MDR/TAP), member 9), whose MDR/TAP sub-family is involved in drug resistance.

Other associations are discussed in the next section. Given the p>5e-08 significance levels of most variants in Table 5.2, when considering a genome-wide panel of markers, validation of top associations is crucial for identifying true associations among false positives.

5.4 Genome-Wide Meta-Analysis of Capecitabine-Related Toxicity in the QUASAR2 and SCP Studies

In addition to the QUASAR2 patients, 233 patients from the SCP study were available for analysis. From this study, only diarrhoea and HFS clinical data were available and patients were selected for genotyping based on high (grade 3/4) or low (grade 0/1) levels of these phenotypes. As described in Methods (Section 2.1) and Chap-

ter 4, these patients were genotyped on Illumina Hap610 arrays and their genotype data were subsequently imputed (IMPUTE2), filtered and analysed (SNPTEST2) according to the same methods used for QUASAR2. Meta-analysis was performed for only diarrhoea and HFS, under the binary (grade 012v34) measure. From SCP, 4,514,766 SNPs were available for analysis after filtering, of which 3,634,937 overlapped with the content available from QUASAR2. Given the disparity in sample size between SCP and QUASAR2, n>900 was set as an inclusion criteria for analysis, with 1,142,840 polymorphisms analysed in QUASAR2-only (n=940) and 3,634,937 variants meta-analysed with QUASAR2 and SCP (n=1,155 for diarrhoea and 1,167 for HFS).

Figure 5.2 shows the additive model GWAS results for all phenotypes and codings of QUASAR2 with or without SCP, as Manhattan and QQ plots. No systematic inflation of the test statistic was observed in any GWAS results. For global toxicity, quantitative codings and the rare phenotypes, results are for QUASAR2-only (top associations discussed in Section 5.3), whereas results for diarrhoea grade 012v34 and HFS grade 012v34 are for a QUASAR2/SCP meta-analysis of the beta and standard error of each SNP (performed in GWAMA). Table 5.3 shows the most significant associations from the diarrhoea and HFS grade 012v34 GWAS meta-analyses – including all SNPs at p<1e-06 and nonsynonymous variants at p<5e-06. Top associations for other phenotypes/codings, from QUASAR2-only, were reported in Table 5.2. All associations with grade 012v34 diarrhoea and HFS from QUASAR2-only (Table 5.2) are also included in Table 5.3 to show the effect of meta-analysis.

For the measure of global toxicity, encompassing the six 5-FU-related phenotypes, no SNP reached genome-wide significance, in either the binary or the quantitative model. The strongest associations were on chromosomes 1 and 18 – the *DPYD* and *TYMS/ENOSF1* associations discussed in Chapter 4. For diarrhoea, five associations at p<1e-06 exist under the binary grade 012v34 model. Two are at loci on

Q2/SCP meta: All SNPs at p<1e-06; nonsynonymous SNPs at p<5e-06

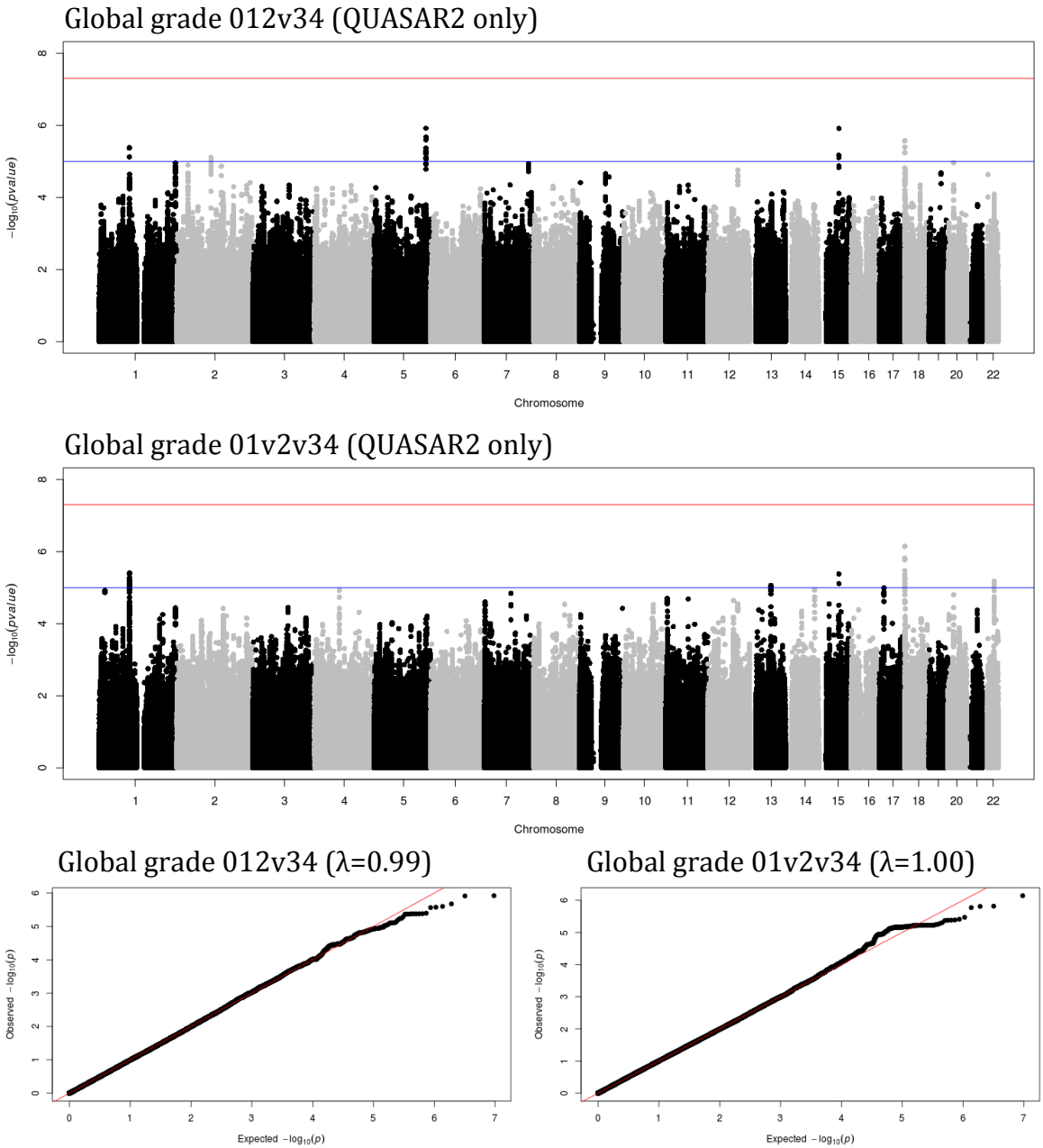
Coding	Toxicity	Model	Variant	Chr	Coord	Ref	Test	OR	LCI	RCI	Pval	MAF	SampleSize	Studies	Q p-value	I-sqrd	Function	Gene
012v34	Diarrhoea	ADD	rs4309833	4	10666766	A	G	0.51	0.39	0.67	7.04E-07	0.41	1155	2	0.06	0.72	intronic	CLNK
							Q2	0.45	0.33	0.60	1.33E-07	0.42	940					
							SCP	0.83	0.47	1.45	5.09E-01	0.39	215					
012v34	Diarrhoea	ADD	rs77987579	4	149618769	A	T	4.78	2.59	8.81	5.38E-07	0.07	940	1	1.00	NA	intergenic	NR3C2 (dist=255097), DCLK2 (dist=1380657)
012v34	Diarrhoea	ADD	rs78285023	7	121488024	T	C	23.35	6.83	79.85	5.22E-07	0.01	940	1	1.00	NA	intergenic	FAM3C (dist=451602), PTPRZ1 (dist=25135)
012v34	Diarrhoea	ADD	rs78160146	10	71716753	C	T	13.35	5.06	35.22	1.69E-07	0.02	1155	2	0.60	0.00	intronic	COL13A1
							Q2	10.61	2.89	38.98	3.76E-04	0.01	940					
							SCP	17.79	4.15	76.26	1.06E-04	0.03	215					
012v34	Diarrhoea	ADD	rs6126708	20	51572585	A	T	6.79	3.18	14.50	7.58E-07	0.04	940	1	1.00	NA	intergenic	ZFP64 (dist=764061), TSH22 (dist=16361)

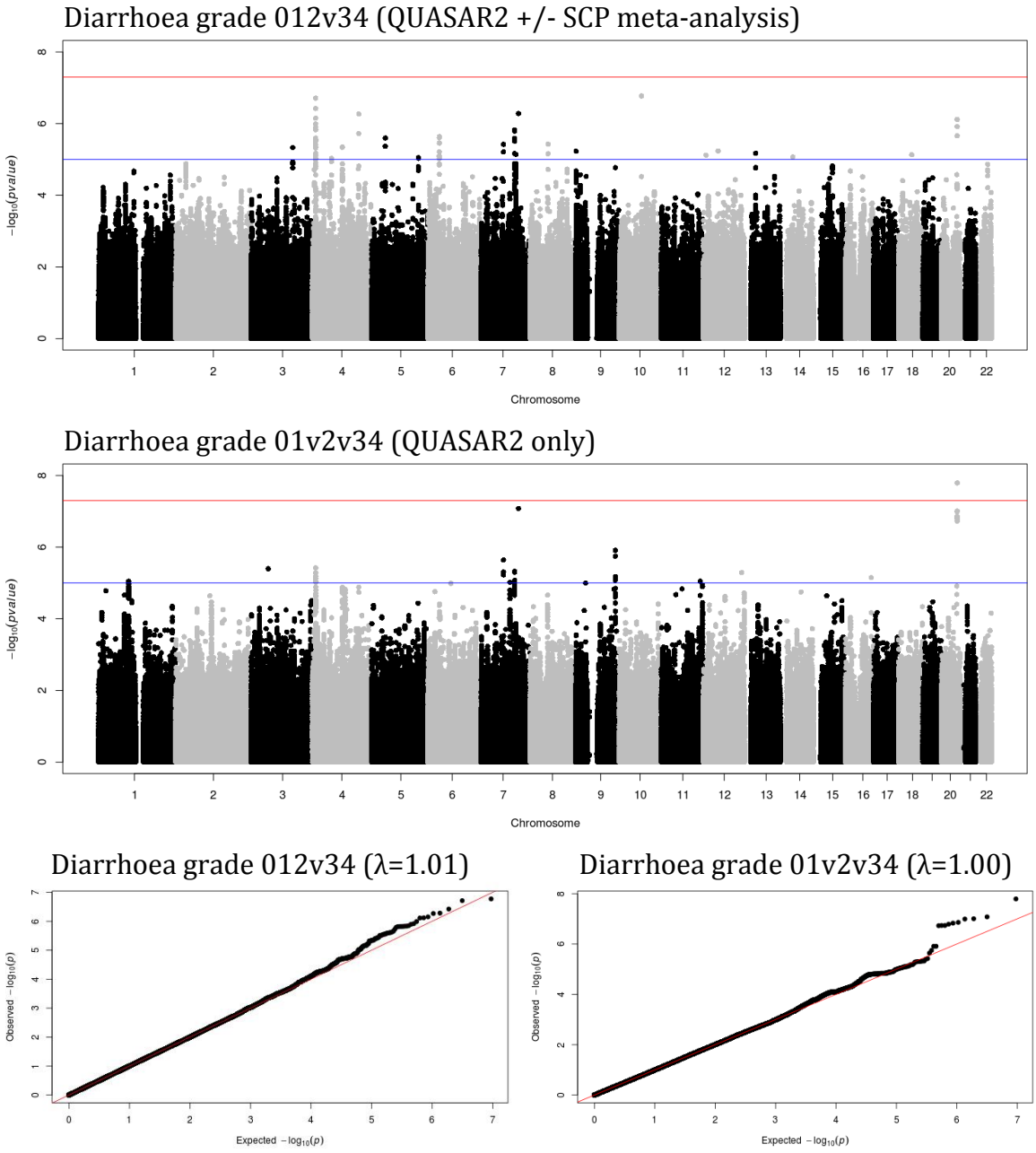
Q2/SCP meta: QUASAR2-only top associations (reported in previous table) after meta-analysis

Coding	Toxicity	Model	Variant	Chr	Coord	Ref	Test	OR	LCI	RCI	Pval	MAF	SampleSize	Studies	Q p-value	I-sqrd	Function	Gene
012v34	Diarrhoea	ADD	rs75153039	4	149609427	A	C	2.78	1.69	4.57	5.96E-05	0.07	1155	2	0.00	0.90	intergenic	NR3C2 (dist=245755), DCLK2 (dist=1389999)
							Q2	4.69	2.58	8.52	4.05E-07	0.07	940					
							SCP	0.84	0.34	2.07	7.02E-01	0.10	215					
012v34	Diarrhoea	ADD	rs7861903	9	128282914	G	A	2.55	1.35	4.80	3.75E-03	0.04	1155	2	0.00	0.95	intronic	MAPKAP1
							Q2	7.33	3.31	16.21	8.86E-07	0.04	940					
							SCP	0.41	0.14	1.16	9.30E-02	0.06	215					
012v34	Diarrhoea	ADD	rs2093152	20	51559324	G	C	6.18	2.91	13.14	2.19E-06	0.03	1155	2	0.00	0.91	intergenic	ZFP64 (dist=750800), TSH22 (dist=29622)
							Q2	10.55	4.66	23.90	1.60E-08	0.03	940					
							SCP	0.30	0.04	2.10	2.25E-01	0.02	215					

Table 5.3: Results from the QUASAR2/SCP GWAS meta-analysis for diarrhoea and HFS grade 012v34 are shown. The top panel includes all associations a p<$1x10^{-6}$ and nonsynonymous SNP associations at p<$5x10^{-6}$. The bottom panel shows the meta-analysis results for those SNPs shown in the QUASAR2-only Table 5.2 that fell above the threshold upon meta-analysis. Given the n>900 inclusion criteria, which allows for SNPs to only be genotyped in QUASAR2, not all SNPs presented were meta-analysed. The rs77987579 SNP was not included in the SCP study analysis, but its perfect proxy rs75153039 (shown here) was meta-analysed and did not replicate. The rs78285023 SNP was also not included in the SCP study analysis and no proxies were identified. The rs6126708 SNP was removed for low SNPTEST info score (0.92) from the SCP study, where its effect was opposite to that observed in QUASAR2 (beta=-1.26, se=0.96, p=0.19); its proxy rs2093152 was among the best associations in the QUASAR2-only analysis but did not replicate, as shown here. In addition, rs2612091/rs2741171 on chr18 and rs12132152/rs76387818 on chr1 were significant, as described in Chapter 4.

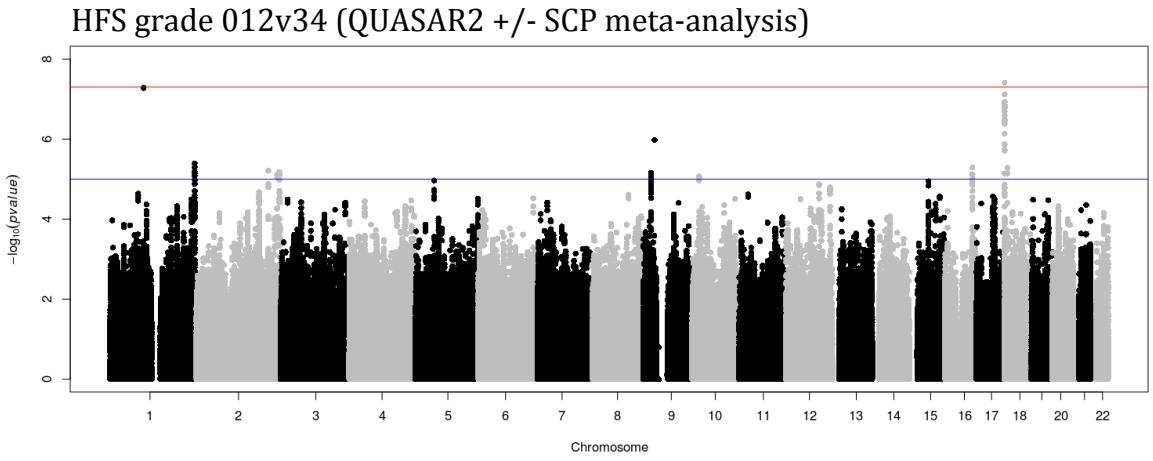
Figure 5.2: Manhattan and QQ plots are shown for the GWAS of global toxicity and each of the six component toxicity phenotypes individually. All analyses include the 940 patients from QUASAR2 for 4.8 million SNPs. In addition, where data were available from the SCP study (diarrhoea and HFS grade 012v34 for 3.6 million SNPs), meta-analysis of SNPs was performed in GWAMA. Lambda values show that no inflation was observed for any of the analyses. Genome-wide significance is marked at $5x10^{-8}$.

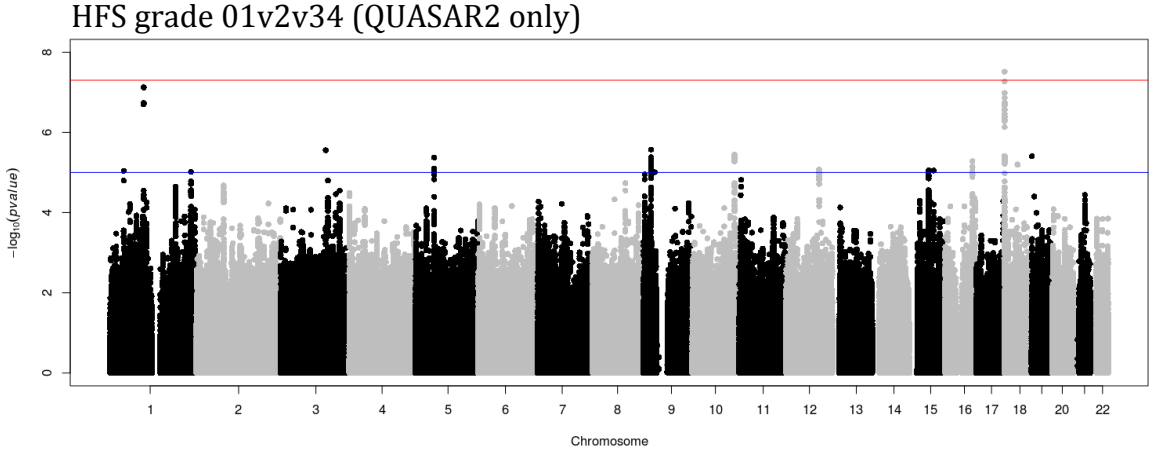


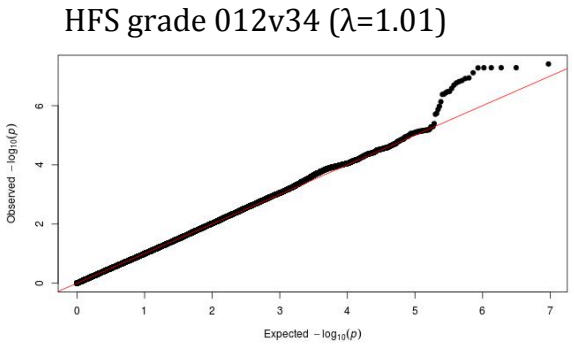


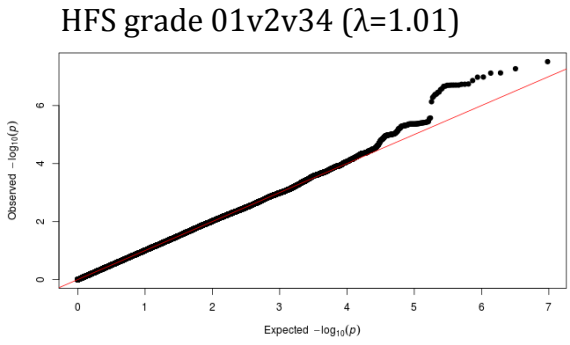
For binary diarrhoea (grade 012v34), SNPs at p<1e-05, where the observed distribution of p-values diverges from the expected distribution, include (top association per chromosomal region listed): rs77987579 - chr4:149618769, rs115172840 - chr5:150780579, rs71631201 - chr5:43975021, rs56738294 - chr6:36954820, rs6457993 - chr6:37517485, rs77124780 - chr7:109142652, rs78285023 - chr7:121488024, rs6960748 - chr7:73302462, rs189640704 - chr8:57770539, SNP9-1801361 - chr9:1811361, rs78160146 - chr10:71716753, rs139853939 - chr12:48124937, rs16861 - chr12:9359244, rs7982372 - chr13:35451150, rs143559477 - chr14:40123575, rs139075675 - chr18:42131925, and rs6126708 - chr20:51572585.

For quantitative diarrhoea (grade 01v2v34), SNPs at p<5e-06 include: rs358880 - chr3:55295766, rs4309833 - chr4:10666766, rs74529775 - chr7:109143587, rs78285023 - chr7:121488024, rs4717829 - chr7:73301666, rs7861903 - chr9:128282914, and rs2093152 - chr20:51559324.



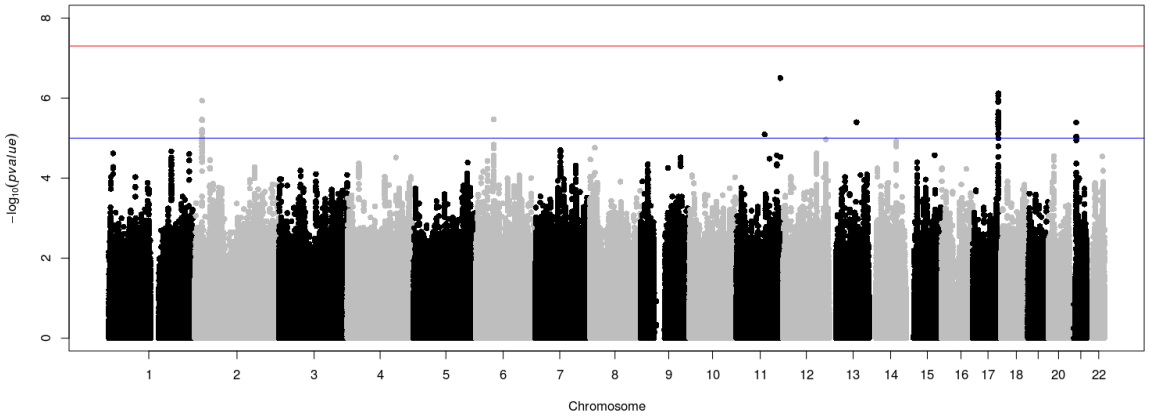




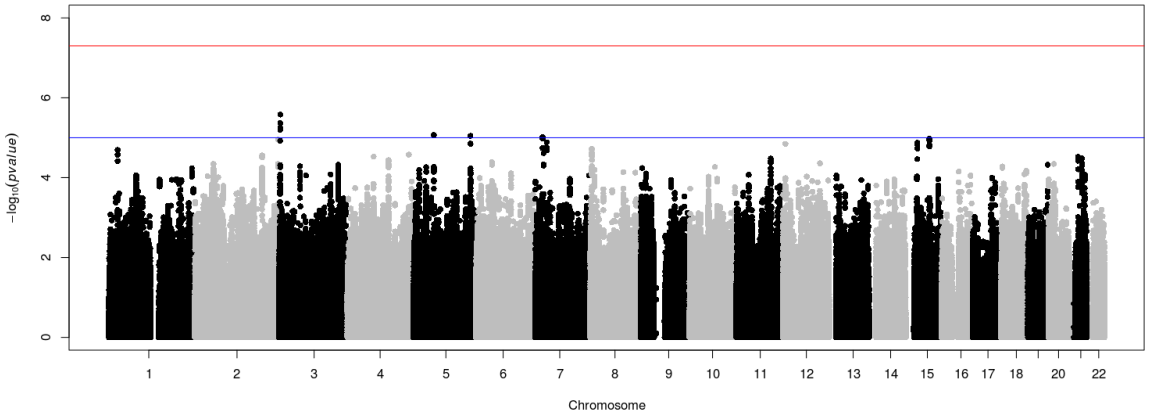


For binary HFS (grade 012v34), SNPs at p<8e-06, where the observed distribution of p-values diverges from the expected distribution, include (top association per chromosomal region listed): rs116448748 - chr1:97490443, rs10754845 - chr1:244807100, rs2284933 - chr2:207999021, rs17220976 - chr9:32741274, rs4381615 - chr16:79185000, rs2847325 - chr18:703442, and SNP18-8784816 - chr18:8794816.
For quantitative HFS (grade 01v2v34), SNPs at p<8e-06 include: rs76387818 - chr1:97539400, rs10804598 - chr3:130087842, rs1023674 - chr5:54265532, rs533644 - chr9:22459115, rs61600621 - chr10:122083762, rs4381615 - chr16:79185000, rs2741171 - chr18:700687, rs73948988 - chr18:37670364, and rs197198 - chr19:1505874.

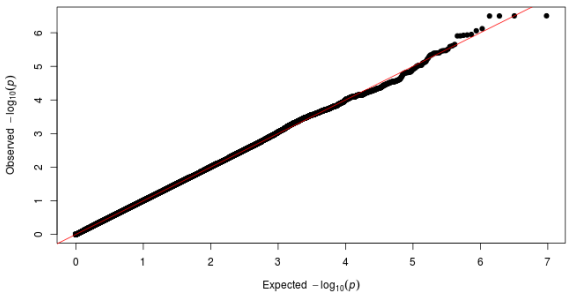
Mucositis grade 0v1234 (QUASAR2 only)



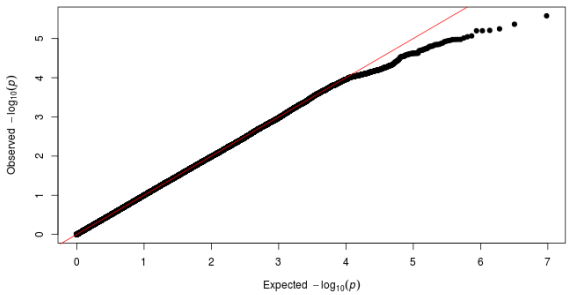
Stomatitis grade 0v1234 (QUASAR2 only)



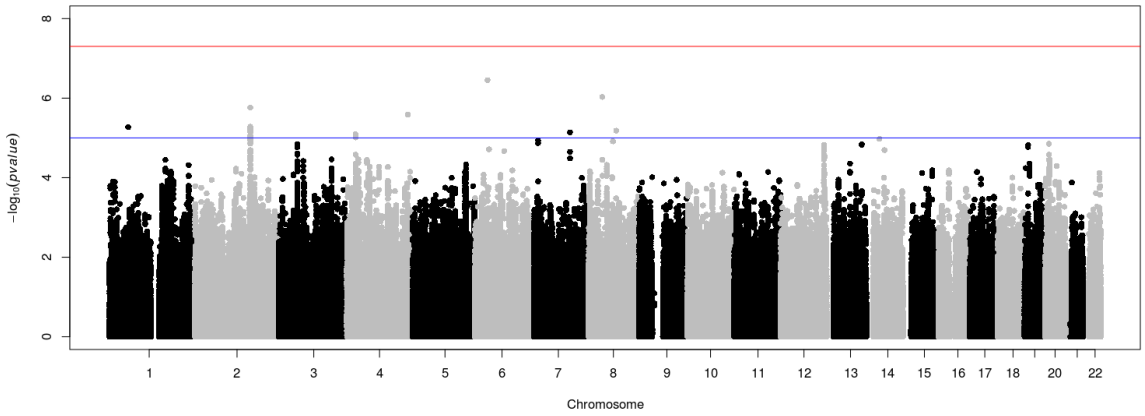
Mucositis grade 0v1234 (λ=1.01)



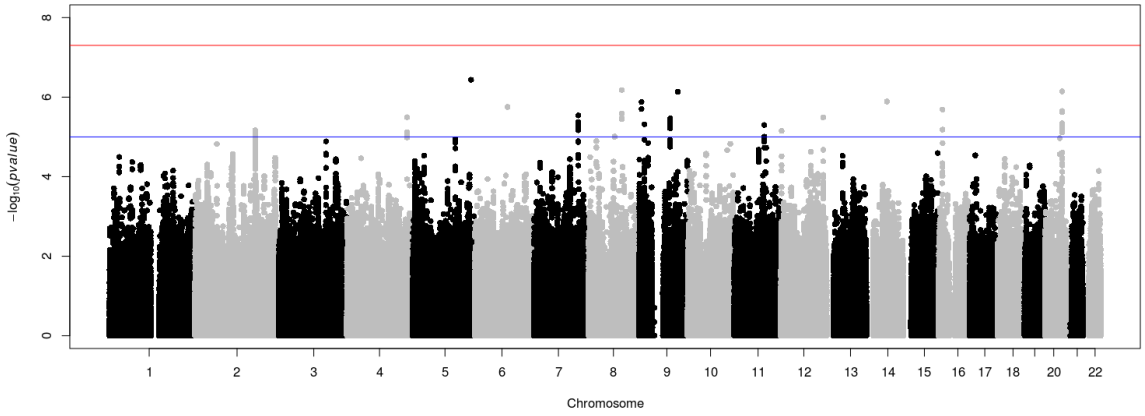
Stomatitis grade 0v1234 (λ=0.99)



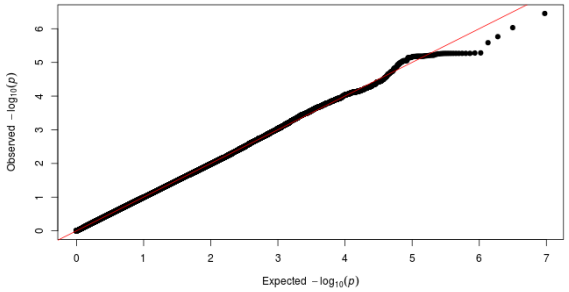
Neutropaenia grade 0v1234 (QUASAR2-only)



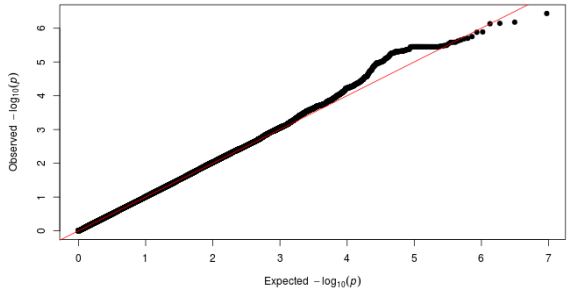
Thrombocytopaenia grade 0v1234 (QUASAR2-only)

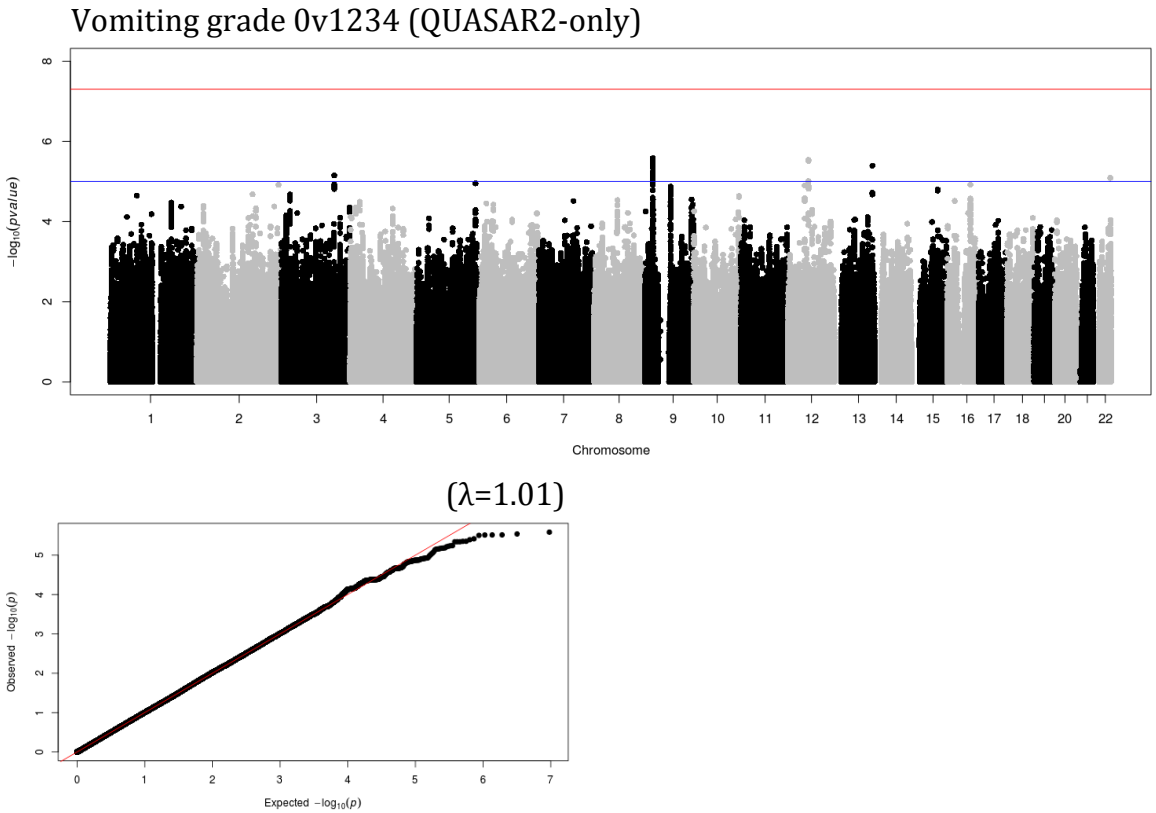


Neutropaenia grade 0v1234 (λ=1.00)



Thrombocytopaenia grade 0v1234 (λ=1.01)





chromosome 4. The first, rs4309833 at chr4:10666766, is a rare SNP intronic to the *CLNK* gene (cytokine-dependent hematopoietic cell linker), which is a modulator of immune system signaling. This variant showed strong evidence of an association in QUASAR2, with little contribution from SCP. The second, rs77987579 at chr4:149618769, was only genotyped in QUASAR2, but its perfect proxy rs75153039 failed to replicate. The rs78285023 (chr7:121488024) SNP is another association included in only QUASAR2; this SNP does not fall near to any gene with a clear rationale for impacting toxicity levels, but is a good candidate for further replication. The rs78160146 SNP association at chromosome 10 was the result of modest significance in both the QUASAR2 and SCP analyses. This rare SNP (MAF=0.02) is intronic to *COL13A1* (collagen, type XIII, alpha 1), where it exists in an area of strong transcriptional activity according to RNA-seq performed as part of ENCODE. It is unclear how this gene would impact toxicity, but the locus is potentially func-

tional. The fifth diarrhoea association, rs6126708 at chr20:51572585, was removed for a low info score in SCP; its proxy rs2093152 did not replicate.

In summary, from the QUASAR2/SCP meta-analysis of grade 012v34 diarrhoea and HFS, for diarrhoea, rs4309833 and rs78160146 remained significant at p<1e-06 upon meta-analysis and rs78285023 showed strong evidence (p<1e-06) in QUASAR2 without showing evidence of poor replication either directly or via a proxy.

5.5 Correlation Among Capecitabine-Related Toxicity Phenotypes in QUASAR2

As discussed in the Candidate Gene Chapter, some SNPs are associated with multiple toxicity phenotypes – as example (Section 4.5.4), the intronic *DPYD* SNP rs7548189 was associated with diarrhoea (p=0.0020), HFS (p=0.0097) and vomiting (p=0.00056), resulting in an association with global 5-FU-related toxicity (p=3.3e-05), and these individual phenotype associations remained when adjusting for the other phenotypes. Although the results of the GWAS presented in this chapter have not identified further variants with significant associations with global toxicity, as additional samples are added to the analysis, it will be important to determine whether SNPs associated with multiple toxicity phenotypes independently influence risk of each phenotype or whether this is the result of inherently correlated phenotypes. In terms of toxicity mechanism, a SNP that is independently associated with multiple toxicity phenotypes might increase the active dose of the drug (ie, a shared mechanism) or act through separate mechanisms to cause each toxicity, whereas a variant that is non-independently associated with two correlated phenotypes might specifically impact one phenotype whose existence causes a second phenotype (eg, neutropaenia causing diarrhoea). To investigate the independence of SNPs associated with multiple toxicity phenotypes, this section explores both the correlation of

toxicity events and the correlation of SNP effect sizes among phenotypes.

Correlation among 5-FU/capecitabine-related toxicity phenotypes was assessed using the Spearman's rank correlation coefficient ("rho"), which tests whether the order of one toxicity phenotype's grade (0, 1, 2, 3, 4) can be predicted with the order of a second phenotype's grade. (It is similar to a Person's correlation coefficient, but allows more flexibility for the (non-)linearity of the two phenotypes' relationship and for outliers.) As seen in Table 5.4, modest correlation exists among some 5-FU/capecitabine-related phenotypes when considering a quantitative measure of the phenotype, but very little correlation exists among phenotypes when considering the binary measure. Specifically, under the quantitative measure, diarrhoea is somewhat correlated with HFS, mucositis, stomatitis, and vomiting, while HFS is also correlated with mucositis and stomatitis. In the binary measure, mucositis and stomatitis are highly correlated, and to a lesser degree so are the highly clinically impactful thrombocytopaenia (platelets) and neutropaenia. As expected, all phenotypes show some correlation with the global measure, according to their incidence rate.

For diarrhoea and HFS – the most common phenotypes and thus the most likely to both show a significant association with a SNP – the modest correlation can be visualised in Figure 5.3, which plots the beta coefficient for an effect on diarrhoea and HFS for the 13,421 SNPs with $p<0.05$ for both phenotypes, among the 4.9 million tested in at least QUASAR2 in the QUASAR2 +/- SCP analysis for either phenotype (10% more than expected by chance – (4,900k)(0.05)(0.05)=12,250). As the phenotypes are largely not correlated, there is not a single positively or negatively sloped line, but rather an "X" shape that is somewhat over-prevalent as a positive correlation (bottom left to upper right line).

The low correlation among phenotypes makes it unlikely that one phenotype induces a secondary phenotype effect; however, to further test this hypothesis, diarrhoea and HFS were tested in multivariate models in which the effect of the SNP on one

	Global	Diarrhoea	Handfoot	Mucositis	Neutropaenia	Platelets	Stomatitis	Vomiting	
Global		0.47	0.74	0.21	0.15	0.09	0.23	0.23	Quantitative
Diarrhoea	0.48		0.15	0.21	0.08	0.05	0.22	0.32	
Handfoot	0.78	0.00		0.20	0.02	0.01	0.23	0.06	
Mucositis	0.14	0.09	0.07		0.06	0.03	0.32	0.09	
Neutropaenia	0.21	0.06	-0.02	0.18		0.25	0.06	0.14	
Platelets	0.09	0.08	0.00	0.30	0.42		0.06	0.03	
Stomatitis	0.15	0.08	0.09	0.69	0.23	0.43		0.14	
Vomiting	0.17	0.14	0.03	0.07	0.04	-0.01	0.06		
	Binary								

Table 5.4: The correlation among the global and individual 5-FU/capecitabine-related toxicity phenotypes was assessed using Spearman's rho, which shows positive and negative correlation. A value of 0 means that no correlation exists among the two phenotypes tested, while a value of 1 means there is a perfect positive correlation. Both binary (grade 012v34) and quantitative (grade 0v1v2v3v4) measures for all phenotypes were assessed.

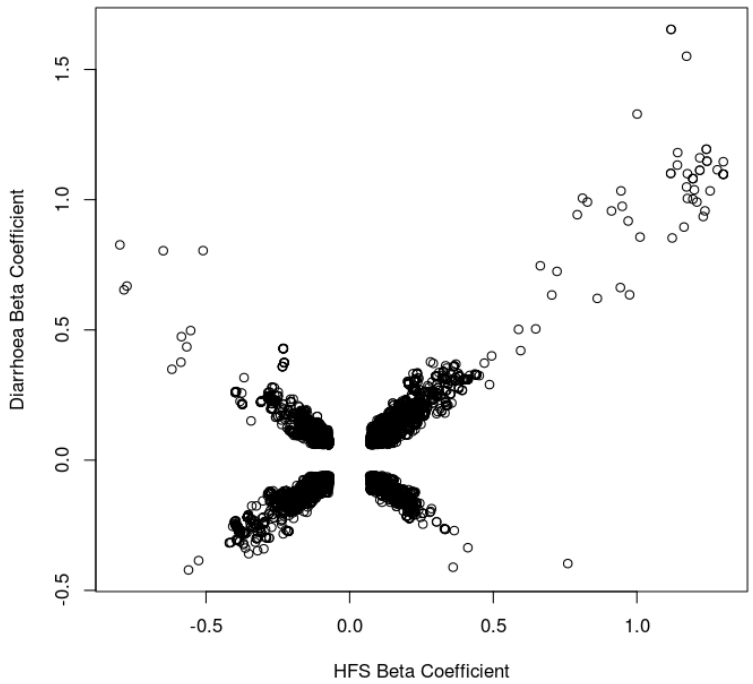


Figure 5.3: The effect size (beta coefficient) of HFS and diarrhoea grade 01v2v34 are plotted for the 13,421 SNPs with p<0.05 for both phenotypes, using the output from the QUASAR2 +/- SCP analysis. Spearman's rho=0.15.

phenotype could be adjusted for the presence of the second phenotype. In other words, is diarrhoea still more likely to be caused by the SNP when we adjust for the incidence of HFS? For this analysis, optimally, the tested SNPs would have a large effect size on both phenotypes. The approximately 50 SNPs from Figure 5.3 with a consistent large beta for both phenotypes, however, were all very rare (MAF<0.001), suggesting that their aligned effect size might be by chance. Instead, among all SNPs, those with p<0.01 and a beta>0.1 or <-0.1 in both quantitative diarrhoea and HFS were extracted and tested in both a standard model (one phenotype, adjusted for age, trial arm and gender) and when adding the other phenotype as a covariate. As shown in Table 5.5, three SNPs met these criteria: rs7548189, rs10745513 and rs67039552, of which the first was previously discussed above in Section 4.5.4. For all three SNPs, significance for each of HFS and diarrhoea is only somewhat decreased when adjusting for the other phenotype, suggesting again that the two phenotypes are not fully dependent.

	P-value	Beta	AIC
rs7548189			
Global ~ SNP	8.79E-06	0.20	
Diarrhoea ~ SNP	2.00E-05	0.16	1898.0
Diarrhoea ~ SNP + HFS	3.67E-05	0.16	1898.8
HFS ~ SNP	1.20E-03	0.15	2223.7
HFS ~ SNP + Diarrhoea	2.24E-03	0.14	2224.5
rs10745513			
Global ~ SNP	5.77E-03	0.28	
Diarrhoea ~ SNP	1.34E-03	0.27	1885.5
Diarrhoea ~ SNP + HFS	2.25E-03	0.26	1885.8
HFS ~ SNP	1.37E-03	0.32	2207.5
HFS ~ SNP + Diarrhoea	2.32E-03	0.30	2207.8
rs67039552			
Global ~ SNP	1.76E-03	-0.13	
Diarrhoea ~ SNP	3.06E-03	-0.10	1888.5
Diarrhoea ~ SNP + HFS	4.55E-03	-0.10	1888.7
HFS ~ SNP	7.70E-03	-0.11	2199.8
HFS ~ SNP + Diarrhoea	0.012	-0.10	2200.0

Table 5.5: For SNPs associated with both diarrhoea and HFS grade 01v2v34 toxicity, the independence of each phenotype's association was tested by multivariate logistic regression. All analyses were adjusted for age, arm and gender. Three SNPs showed a consistent positive or negative effect for HFS and diarrhoea; in all instances, each association was largely independent.

The implication of the mostly low correlation rates among phenotypes is that a SNP associated with multiple phenotypes likely acts either via a single mechanism shared by both toxicities (eg, increased active drug dose) or acts on multiple pathways that each independently affect one toxicity, but is unlikely to act specifically on one phenotype that itself influences/causes a second phenotype. Given the assumptions of Mendelian Randomization, whereby inheritance of one trait is unlikely to be coded by the same SNP as another trait, it seems probable that SNPs associated with multiple phenotypes do so as a result of a shared mechanism. As additional SNPs with strong associations with multiple phenotypes are discovered, it will be interesting to identify whether any have a clear rationale for a single specific phenotype, in which case a secondary effect from one phenotype to another phenotype might be identified.

5.6 Discussion

This chapter reported the genome-wide results of the QUASAR2-based association study between germline polymorphisms and capecitabine-related toxicity. From the QUASAR2 study, over 4,700k directly genotyped or imputed SNPs were tested for an association with global toxicity and with each of the six phenotypes that comprise that global measure, as were over 230k rare and largely coding exome array variants. While no variants reached genome-wide significance for the global measure, rs2093152, an intergenic SNP between two zinc finger genes, fell below this threshold for grade 012v34 diarrhoea in QUASAR2 and additional SNPs were near genes with a plausible mechanism for impacting toxicity. The rs2093152 SNP, however, failed to replicate in a meta-analysis with the SCP study, for which no variants reached genome-wide significance for grade 012v34 diarrhoea or HFS. From the QUASAR2-only analysis of all phenotypes and QUASAR2/SCP meta-analysis, multiple SNPs exist at the $p<1e-06$ threshold for multiple phenotypes and these variants would benefit greatly from

further replication efforts. Although no new toxicity variants have been definitively identified, the significance levels achieved in this study surpass those of previous chemotherapy toxicity GWAS analyses.

The estimation of heritability was less successful, likely the result of sample size. Commonly 2,000 or more patients are included in such analyses, whereas approximately 1,000 were available for this analysis. If additional samples from QUASAR2 or another large UK-based capecitabine study become available, a more accurate prediction of the heritable component of toxicity should be possible.

Using QUASAR2 alone, an initial assessment of correlation among phenotypes was performed. While most phenotypes are only marginally correlated – particularly when considering them as a binary measure – the modest correlation between diarrhoea and HFS was detectable when adjusting associations with one phenotype for events in the other. However, for three SNPs with modest effects in both phenotypes, associations appear largely independent.

In summary, while a short list of variants has been identified by GWAS, further replication will be necessary to determine true associations with certainty.

Chapter 6

Germline Markers of Prognosis in CRC

6.1 Introduction

Little is known about germline genetic markers of colorectal cancer prognosis, as reviewed in Introduction. Current treatment decisions are based on clinical factors such as histopathological disease stage and patient age. Somatic (ie, tumour) DNA changes further help sub-classify some patients, using factors such as chromosomal instability (CIN), microsatellite instability (MSI) and single nucleotide variants in *KRAS* and *BRAF*. However, there exists sparse reporting of germline genetic markers of CRC prognosis. Genome-wide, no such study has been published, the closest being a previous attempt at analysing the VICTOR dataset by another DPhil student in the Tomlinson lab, Dr Axel Walther, which formed the basis of an ASCO poster in 2010. Of the top hits from this analysis, one hit near the cell motility gene *ACTR2* reached the genome-wide threshold upon meta-analysis in four replication sets. However, this result awaits a further replication set.

Since that work, extended patient follow-up data have become available and it has

come to light that the VICTOR trial possesses a severe left truncation that might have reduced statistical power – the focus of this chapter. Subsequently, multiple groups have hypothesised that the known CRC susceptibility SNPs might impact survival, and one group has reported a positive finding for a SNP near *SMAD7* (153). Given that data collection for that meta-analysis was not systematic, the result is less strong; however, the variant's effect was in the same direction (though not significant) in previous results of larger sample size not included in the meta-analysis (192). Beyond this SNP, the literature remains limited in the study of CRC prognosis, and making use of the VICTOR dataset for a genome-wide analysis would provide a useful addition to the field.

The VICTOR dataset contains an unusually large left truncation in the form of a sampling bias of patient blood collection for genetic study. Truncation and censoring occur when there are incomplete data for a patient. In most clinical trials, modest amounts of two forms of data loss occur:

Right censoring: observed event (eg, death) occurs after censoring.

Left truncation: observed event occurs before trial recruitment.

Right censoring is common to all trials, as the end date of patient follow-up is often pre-specified, both to avoid favorable selection of outcome and to provide results in a reasonable timeframe. Thus, while all patients die *one day*, 5-year overall survival and 3-year disease-free survival are common points at which follow-up ends (ie, when all remaining patients are censored). Large left truncation is less common in most trials, though it cannot be directly measured. In CRC, left truncation typically takes the form of patients who *would have been* recruited to a study had they not died in the period between their diagnosis/surgery (when they become eligible for the study) and when trial recruitment/enrollment began. This recruitment/enrollment date typically marks the start of a patient's "at risk" period, namely:

At risk period: span of time when an event could be recorded for a

patient, usually from recruitment to censoring.

These censoring, truncation and at risk period descriptions are all in regards to a study of an *intervention* and its effect on survival. In regards to a study of the effect of a germline *polymorphism* on survival, left truncation can be enormously exacerbated, as is observed in the VICTOR trial. This is because the at risk period for a patient does not begin at trial recruitment, but rather at blood collection, when we obtain his or her germline genetic data – a phenomenon known as "survival bias". Thus in a germline genetic analysis of VICTOR prognosis, patients are only "at risk" once they have both been recruited to the study (meaning we could observe their death) and had their blood collected (meaning we can identify their genotype). This sampling bias manifests as a reduction to statistical power through missed events and a possible violation of the proportional hazards assumption integral to the Cox proportional hazards method, meaning that the creation of well-powered and non-biased results requires careful model selection.

6.2 Aim and Overview

This chapter applies various methods to investigate germline markers of CRC prognosis. First, it defines the left truncation in VICTOR. It then applies standard techniques to an investigation in this trial, using the Cox proportional hazards model with selected clinical covariates. Next, a series of non-standard analyses are performed with the primary aim of providing validation for the direction of effect observed in the Cox model – first the problem is flipped on its head, using the blood collection lag as an outcome variable instead of as a covariate, both in VICTOR patients and then in patients from the CORGI consortium of colorectal cancer patients; next, it looks to consider VICTOR patients on the whole as cases, given the extent of its truncation, and compares allele frequencies with patients from the QUASAR2 trial,

in which blood was collected without lag. In addition, each of these models is briefly analysed as a GWAS and their top associations are reported. Finally, the chapter merges these various approaches into a novel "full" survival model that incorporates data from both VICTOR and QUASAR2. Results from these approaches are then compared with the aim of identifying SNPs of interest based on consistent results in multiple analyses.

6.3 Outcome Measures

In studying CRC prognosis, this thesis uses two outcome measures: overall survival (OS) and recurrence-free survival (RFS), as indicated in the VICTOR trial protocol. However, because the length of time from surgery to randomisation was noted to differ substantially based on whether or not a patient was treated with chemotherapy, time-to-event endpoints are measured from surgery, not randomisation. The definitions followed for OS and RFS are provided here; in addition disease-free survival (DFS), a commonly used measure, is defined:

OS: time from surgery to death from any cause.

RFS: time from surgery to either recurrence or death from CRC (ie, CRC-specific survival).

DFS: time from surgery to either recurrence or death from any cause.

Of the 918 VICTOR patients described in this chapter, there are 108 OS events and 189 RFS events, the intersect of which is 91 patients who died from CRC. DFS is the union of OS and RFS, for which there are 206 events. While DFS maximises recorded events, it adds only 9% to RFS. For this chapter, OS and RFS are investigated as, in the case of a finding, their biological implications are more clear, namely that an RFS biomarker is specific to one's propensity to progress with CRC, while improved OS is the true ultimate clinical goal. Using DFS, the deaths unrelated to CRC may act as

a confounder and limit power. In the literature, DFS is mostly useful as a surrogate for OS when follow-up time is limited to 3 years, whereas VICTOR follow-up is 7 years and, as shown later in this chapter, the recurrence curve flattens out prior to end of follow-up.

6.4 VICTOR Dataset Summary and Truncation

6.4.1 VICTOR Clinical and Molecular Data

The VICTOR trial of the COX-2 inhibitor rofecoxib (Vioxx®, Merck) provided data for the primary analysis of CRC prognosis. As described in Methods, of the 2,434 stage II/III CRC patients entered into the study, bloods were collected for 969 patients, of whom 934 were available for analysis following genotyping quality control. Clinical data for these patients were provided by Sue Dutton of the Centre for Statistics in Medicine. Of the remaining 934 patients, I removed a further 13 patients because their clinical data suggested a (impossible) negative overall survival genetic analysis "at risk" period, meaning that their date of death was listed to be as much as two years before their date of blood collection. I removed a further 3 patients whose date of overall survival and/or recurrence-free survival did not include a year, leaving 918 patients for genetic analysis out of a total of 2,383 patients with or without blood collection but who were not removed for quality control measures. Of these patients, Dr Enric Domingo identified that 5 patients were in fact stage I CRC patients, based on a manual review of the pathology reports; these patients were left in the analysis. Demographics for the 918-patient set were as follows: 595 males, 323 females; 600 colon, 318 rectal; 464 stage I/II, 454 stage III; 572 receiving adjuvant chemotherapy, 346 receiving surgery only. Of these, 108 experienced an overall survival event and 189 experienced a recurrence-free survival event (91 patients included in both measures), during the 7-year maximum follow-up period for which data were available.

6.4.2 Left Truncation of the VICTOR Dataset

Blood collection for the VICTOR trial only *began* seven months after the end of patient recruitment. As shown in Figure 6.1, in order to be "at risk" for overall survival following diagnosis, patients must have survived through periods of surgery, chemotherapy and the trial duration prior to when blood collection was initiated – and then be among the subset of living patients for whom blood collection occurred. The scenario is similar for a study of recurrence, except that an additional complication exists in that patients can have a recurrence, survive this recurrence, and then subsequently have blood collected and be entered into the genetic study.

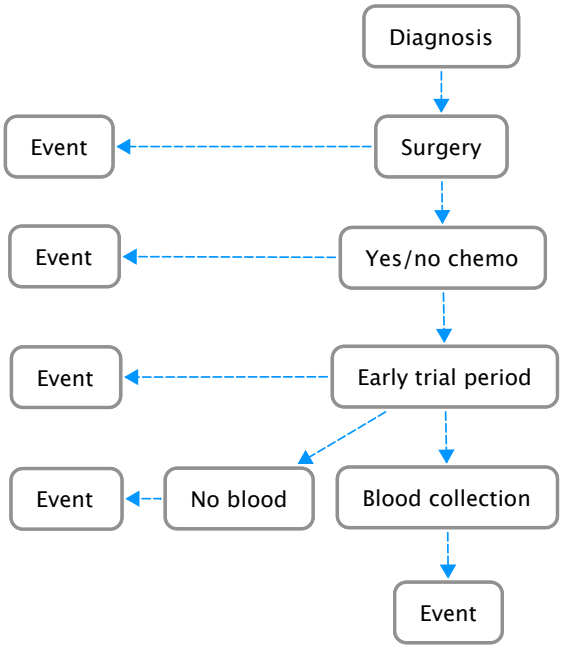


Figure 6.1: Patients in the VICTOR trial only became "at risk" in terms of a genetic analysis of prognosis if they survived until blood collection began midway through the trial and then were among the subset of patients from whom blood was taken. Events at left are all missed in a genetic study, while the event at the bottom is both observed and is complemented by genetic data.

Figure 6.2 compares the left truncation in VICTOR when testing an intervention and when testing germline genetics, in regards to overall survival. When considering all 2,383 patients eligible for a survival analysis together, 13 distinct censor-

ing/truncation patient subgroups exist. First is an unknown number of left truncated patients who theoretically would have been entered into the study, except that they died between diagnosis and recruitment start. All patients in the remaining 12 subgroups were successfully recruited to the study following diagnosis (ie, the 2,383 patients eligible for a survival study). Enrollment of a trial does not occur at one time point, but rather occurs over time (April 2002 to September 2004). Thus, while half of these dozen patient subgroups had surgery before April 2002 [2-7] – meaning they were at increased likelihood of dying prior to enrollment and thus being left truncated from any study of survival – the vast majority of patients had their surgery after April 2002 [8-13] and thus were enrolled quickly. Focusing just on this second set of six patient subgroups (n=2,137), we see that following enrollment, patients either died (n=415) or were censored (n=1,722). Some of these patients leave the study from death or censoring prior to when blood collection began in April 2005 (n=172), though most remained in the study during some portion of the blood collection period. However, only a small proportion of patients available for blood collection did in fact provide a sample; across all 2,383 patients, of the 485 total patient deaths that occurred prior to censoring [2-4,8-10], bloods were collected from only 108 patients [3,9]. In a genetic study of VICTOR, many patients never become "at risk".

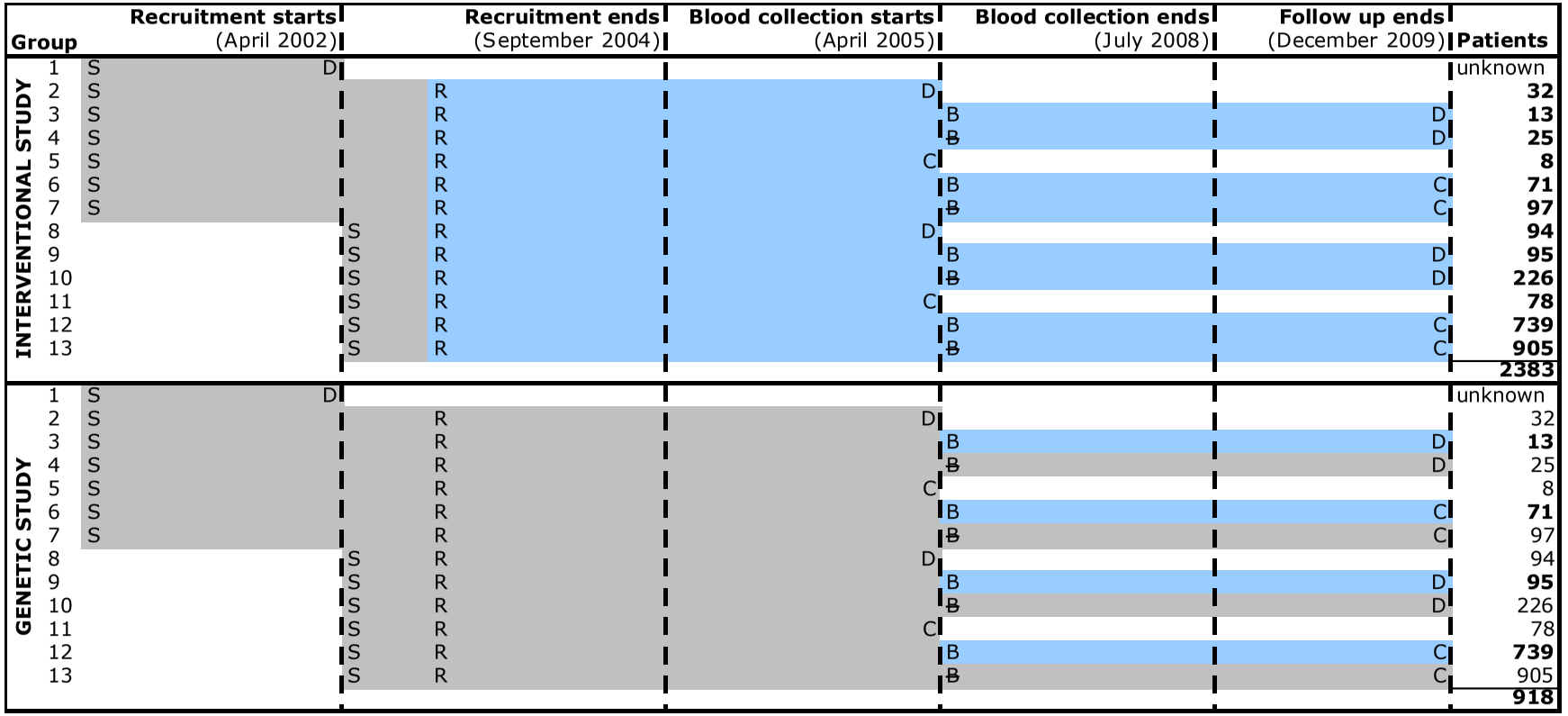


Figure 6.2: Various forms of censoring and truncation exist in the VICTOR trial, and these are exacerbated during a genetic study. TOP: from the perspective of an *intervention-effect* study, patients become "at risk" upon recruitment/enrollment into the trial and remain at risk until death or censoring. For this type of study, 2,383 patients become at risk. BOTTOM: however, from the perspective of a *genetic-effect* study, patients only become "at risk" upon enrollment and subsequent blood collection, both of which must occur prior to death. In VICTOR, blood collection began 3 years after recruitment began (7 months after recruitment ended), and continued for 3 years and 3 months. As a result, a left truncation exists in which only 918 patients become at risk, of whom 108 die prior to censoring. These patients have overall survival periods of up to 7 years, but are only able to be observed as an event in a genetic study (ie, "at risk") for a maximum of 4 years, limiting statistical power. Grey=not yet at risk, blue=at risk; S=surgery, R=recruited, D=death, C=censoring, B=blood collection, crossed B=no blood collection.

6.4.3 Blood Collection Lag and VICTOR Prognosis

The blood collection lag period can be observed when viewing Kaplan-Meier survival curves stratified by whether or not a patient's blood was collected, as shown in Figure 6.3. Whether measuring from surgery or from randomisation, OS and RFS are significantly different between the group of 918 patients from whom blood was collected and the group of 1,465 patients from whom no blood was collected. Outside of a genetic analysis, this blood collection makes no difference, and we observe in the bottom row of this figure that 5-year overall survival for the combined 2,383-patient group is approximately 80%; measured from randomisation, this value is close to 75%, which is in line with established early stage II and early stage III survival in CRC, meaning that patient recruitment in VICTOR shows no evidence of bias.

However, in a genetic analysis of CRC, in which bloods are necessary to obtain germline genotypes and thus only those patients from whom bloods are received become "at risk" in the statistical sense, it is evident that the blood collection lag results in a substantial loss of statistical power. Further, has there been bias, whereby events are not just missed but rather association results are distorted? The survival curves help answer this question. While the early OS and RFS curves appear drastically different between the blood and no-blood groups, they become parallel in later years. This suggests that while early events are missed in the bloods group (because patients must survive until blood collection began – a minimum of 7 months for patients enrolled at the end of the trial, but often much longer), the baseline survival function is not inherently different in the two groups. This is good, because it suggests that blood collection is random among patients who remain alive at a given time from surgery, whereas non-parallel curves would suggest a structural difference between the two groups. This notion is further discussed in Section 6.4.4, where the proportionality assumption is formally tested. From a biologic standpoint, it seems unlikely that a variant would be associated with *only* events in the first year or *only* events in the

later years, and thus while we reduce statistical power through a loss of total events observed, we would expect variants to provide a proportional effect over time (as is assumed in the Cox proportional hazards model) and thus produce the appropriate effect size if sample size and total event rate are sufficient so as to detect it.

As an initial implication, though, these survival curves might imply that the analysis of RFS is statistically more powerful; this model is indeed tested, but as noted later, various complications arise in more intricate analyses of RFS, namely that blood collection can take place following a RFS event (whereas no blood collection occurs following death).

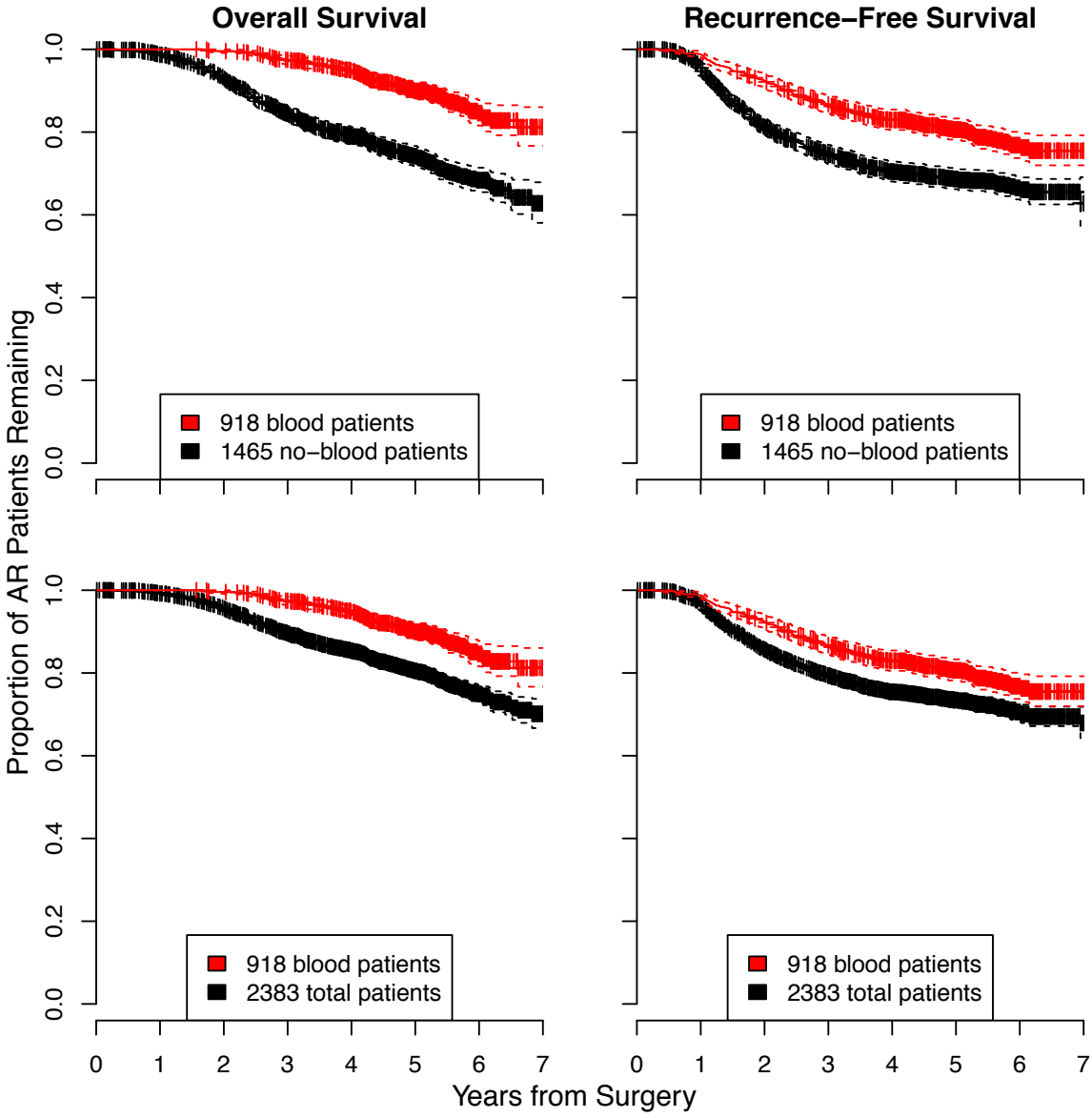


Figure 6.3: As a result of the lag in blood collection, overall and recurrence-free survival times differ substantially between the group of 918 patients from whom bloods were taken and the remaining 1,465 patients from whom no blood was taken. Namely, as all patients in the blood group survived for at least the 7 month period between when the final patient was recruited and the first blood was collected, the blood group is missing all early events. Thereafter, event rates begin to even out between the blood and no-blood group, and by year three the curves are largely parallel (suggesting an accurate hazard ratio), earlier for RFS. This pattern is observed whether comparing the blood and no-blood groups (TOP) or when comparing the blood group to all VICTOR patients (ie, blood group counted twice) – a normal survival curve (BOTTOM). Dotted lines are 95% CIs. Significance: $p<0.0001$ for all (clockwise from top left, chi-squared statistics of 92, 39, 18, and 43). Results remain similarly significant when measuring survival from randomisation instead of from surgery. AR=at risk.

6.4.4 Covariate Selection

Covariate selection for Cox proportional hazards survival analyses was performed with both forward and reverse stepwise removal of non-significant variables in Cox models conditioned on blood collection lag period, testing the clinical variables age (continuous), disease stage (I/II vs III), gender, site (colon vs rectum vs rectosigmoid junction), chemotherapy use (yes or no), and radiotherapy use (yes or no). In addition, the collection blood lag period (equivalent to the difference between a patient's survival date and start of their at risk period) was tested as a continuous variable. This process was repeated for both OS and RFS. As shown in Table 6.1, age, disease stage and gender were selected for inclusion in the overall survival model, while disease stage and radiotherapy use were selected for inclusion in the recurrence-free survival model. For OS, while the blood lag period was not significant in the univariate model, and thus not brought onward in forward stepwise covariate selection, the lag period was significant when adjusted for the other significant covariates, as identified by reverse stepwise covariate selection; including lag also maximised r^2 relative to including just age, stage and gender for OS.

In addition to showing significant effects on survival, these covariates do not significantly differ from demonstrating proportional hazards over time, as identified by Schoenfeld residuals, tested for both OS and DFS for included covariates. Schoenfeld residual p-values are shown in Table 6.1, while plots of covariates for overall survival are shown in Figure 6.4. We see that none of the clinical covariates significantly vary in effect over time, suggesting that the Cox proportional hazards model is valid. This is further supported by the non-significant "global" measure of each model's fit as a whole. Of note, however, for both OS and RFS, the Schoenfeld residual for the blood collection lag period approached significance when not conditioning on lag period (OS: p=0.060, RFS: p=0.080); this potential for bias is one reason why the baseline hazard is fully modeled in the "full" model described in Section 6.10.

	Overall Survival				Recurrence-Free Survival			
Variable (reference)	**Univariate**	**Multivariate**	**OR**	**Schoenfeld**	**Univariate**	**Multivariate**	**OR**	**Schoenfeld**
Age (continuous)	0.019	0.016	1.03	0.56	0.16			
Stage (II)	0.00081	1.3E-04	2.23	0.65	2.4E-06	3.8E-06	2.87	0.18
Gender (male)	0.028	0.031	1.61	0.62	0.24			
Site (colon)	0.97				0.12			
Chemo (none)	0.48				0.29			
Radiotherapy (none)	0.42				6.3E-03	0.013	1.41	0.93
Lag (continuous)	0.13	0.039	0.75	0.51	0.15			
GLOBAL – lag included				0.83				0.46
GLOBAL – lag excluded				0.83				0.39

Table 6.1: Data for various clinical covariates were available for the analysis of overall survival and recurrence-free survival in VICTOR. The covariates included in each final Cox proportional hazards model were selected based on forward and reverse stepwise removal of non-significant covariates. Prior to final inclusion in the model, Schoenfeld residuals were used to test whether proportionality holds, both for individual covariates and for the model as a whole ("global"). In addition to clinical variables, the blood collection lag period (measured from surgery to collection, in years) was tested as a covariate. The same clinical covariates remained significant irrespective of inclusion of lag as a covariate, though significance levels did change and lag period was only significant for OS and when adjusted for the clinical covarites. Values shown are either p-values or odds ratios (OR).

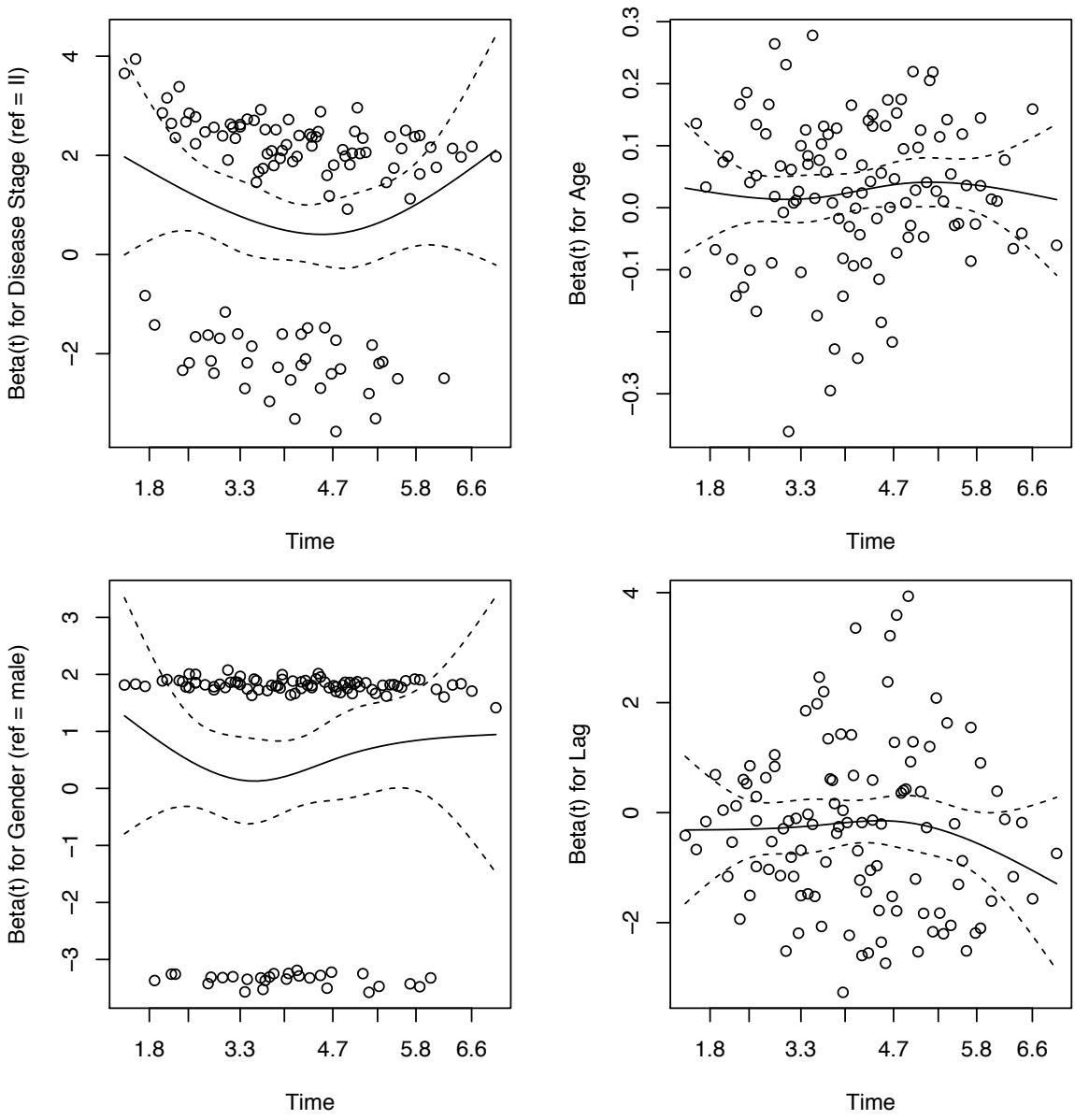


Figure 6.4: Schoenfeld residuals for the covariates included in the analysis of overall survival in VICTOR, showing the effect (beta) of each covariate on a per-event (ie, death/censoring) basis over time. For each covariate, odds ratios (the exponential of the beta) and significance levels across the full time period are shown in Table 6.1, as is the Schoenfeld significance level (p-value) testing the null hypothesis that covariate effect is consistent over time. Covariates with non-significant Schoenfeld p-values – observed as horizontal lines with slope zero – conform to the proportional hazards assumption. Time is in years from surgery.

6.5 Heritability of CRC Prognosis

An estimate of the heritability of CRC prognosis was performed using LDAK software (183) to calculate a weighted kinship matrix of SNP independence and restricted maximum likelihood (REML) analysis using GCTA software (218). While optimally this analysis might take time-to-event and censoring data (as well as interval time) as outcome variables, as occurs in standard survival analyses, the GCTA software is limited to a single binary or quantitative outcome variable. Given that time-to-event data are meaningless when not joined by an indicator variable for whether patients left the study due to death/disease or censoring, heritability had to instead be estimated using binary variables that identified whether overall survival or recurrence-free survival events occurred within the 7-year follow-up period.

Incorporated in the analysis were the covariates described in Section 6.4.4. Although the LDAK portion of the analysis accounts for linkage disequilibrium among included SNPs, population structure differences among cases and controls can cause an inaccurate prediction (28). As such, a principal components analysis (PCA) was performed and the first 10 PCs were included as covariates. As shown in Figure 6.5, the 918 VICTOR patients used throughout this chapter cluster very tightly when plotting the first and second eigenvector in an X/Y space, suggesting low risk of population structure differences resulting in spurious heritability estimates or (later) association tests. As discussed in this chapter, the lag period between patient surgery/recruitment and blood collection (for genetic analysis) can be significant in VICTOR, and thus lag was also used as a covariate.

As shown in Table 6.2, none of the heritability analyses identified a heritable component for CRC prognosis – while the point estimate for the proportion of variance explained by genetics was as high as 47%, the confidence intervals were enormous, ranging from 0% to 100% heritability for all estimates. Thus, rather than suggesting that CRC prognosis is not a heritable phenotype, it is likely that the analysis method,

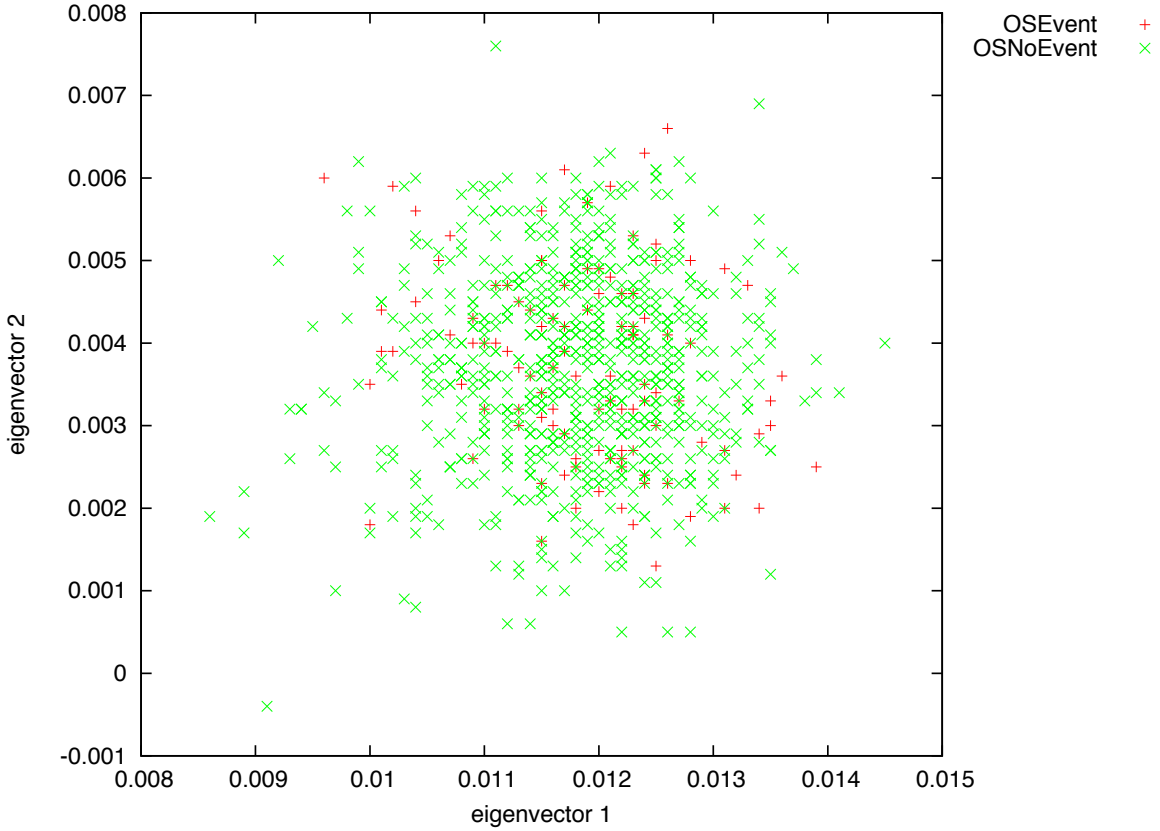


Figure 6.5: The first and second eigenvector of the principal components analysis of the 918 VICTOR patients included in the various analyses of this chapter are shown in an X/Y plot. 270 HapMap population controls (not shown) were used when calculating the coordinates of samples, which cluster with the CEU HapMap samples. VICTOR overall survival (OS) controls are shown in green Xs, while cases are shown in red +s; cases and controls cluster very tightly, suggesting that population structure should cause minimal bias to results.

sample size and/or sample ascertainment method limited power. It is thus interesting to consider what factors might reduce statistical power within VICTOR and how the data might still be used to identify genetic markers of prognosis.

		Variance	SE	Variance	SE	Variance	SE
Overall Survival	**Covariates**	Age, stage, gender		Age, stage, gender, lag		None	
	Vg	4.58%	0.045	4.78%	0.045	3.09%	0.046
	Ve	5.65%	0.045	5.38%	0.045	7.30%	0.046
	Vp	10.22%	0.005	10.16%	0.005	10.39%	0.005
	Vg/Vp	**44.77%**	0.442	**47.03%**	0.441	**29.72%**	0.439
	p-value		0.2		0.1		0.2
Recurrence-Free Survival	**Covariates**	Stage, radiotherapy		Stage, radiotherapy, lag		None	
	Vg	1.39%	0.067	1.77%	0.067	2.51%	0.070
	Ve	14.21%	0.067	13.81%	0.067	13.85%	0.071
	Vp	15.60%	0.007	15.58%	0.007	16.36%	0.008
	Vg/Vp	**8.89%**	0.430	**11.33%**	0.431	**15.35%**	0.430
	p-value		0.4		0.4		0.4

Table 6.2: Binary outcome variables for overall survival (OS) and recurrence-free survival (RFS) were tested for a heritable component using the LDAK/GCTA method. The Vg/Vp value is an estimate of the proportion of phenotypic variance that is attributable to genetic variability. Estimates have been transformed to percentages, whereas standard errors (SE) of these estimates have not. Thus, assuming a normal distribution, the confidence of our estimate would be the raw estimate +/- 1.96*SE, which produces a 95% confidence interval from 0% to 100% heritability for all estimates. Vg=genetic variance estimate, Ve=residual variance (environmental), Vp=phenotypic variance, Vg/Vp=ratio of genetic variance to phenotypic variance. P-values are a log-likelihood ratio on 1 degree of freedom.

6.6 Prognosis in VICTOR by Cox Proportional Hazards Method

The Cox proportional hazards model is the gold standard approach to survival analysis when covariates need to be included in the model. Whereas the Kaplan-Meier method allows one to visualise survival curves for two groups, each further covariate requires stratifying the analysis into additional curves prior to performing logrank tests, reducing statistical power. Thus, upon investigating the need for covariates (Section 6.4.4), the Cox proportional hazards model was selected for genome-wide survival analysis in VICTOR.

The Cox proportional hazards model was conditioned on blood lag period and applied in four ways: with and without applying the blood lag period as a covariate,

for each of OS and RFS. For each analysis, 316,705 autosomal SNPs from Illumina Hap300 arrays were analysed (following SNP and patient genotyping quality control performed by Dr Claire Palles prior to my arrival), for the 918 VICTOR patients from whom blood was collected and who passed my clinical data quality control measures. In addition to the blood lag period, additional clinical covariates were as specified in Table 6.1.

Figure 6.6 shows the Manhattan plot results of these analyses, while Figure 6.7 shows the QQ plots. Viewing both sets of results, it is clear that including the blood lag period as a covariate makes minimal difference to the results; without conditioning on this lag period, we would expect this "minimum survival" period to be far more impactful. Considering the OS results (top row), we observe that two SNPs are below p=1e-06: rs4934251 on chromsome 10 (HR=2.3, p=1.3e-07) and rs9514816 on chromosome 13 (HR=2.2, p=3.6e-07). The rs9514816 variant is 77kb upstream of *LIG4*, a DNA ligase needed for double stranded break repair, while rs4934251 is intronic in the *LDB3* gene, a protein involved in cytoskeleton assembly. Neither variant has been noted as clinically relevant previously, though *LDB3* has been implicated in certain cardiac conditions and mutations in *LIG4* can cause LIG4 Syndrome and have been associated with decreased survival in breast cancer (74). Replication for the *LIG4* SNP failed in previous work by Dr Axel Walther, while replication has not been attempted for the *LDB3* SNP. The RFS results (bottom row) were less strong, with no SNPs falling below p=1e-06.

Concordance among OS and RFS results was modest, with r^2 of 0.15 among all SNPs, as shown in Figure 6.8. Seven SNPs had a beta greater than 1 (HR>2.7) in both OS and RFS (rs10849969, rs10850016, rs11066209, rs11066226, rs2301757, rs4767944, and rs6887391), though each was significant at p<0.05 in only either OS or RFS. 26 SNPs had a beta less than -1 (HR<0.37) in both OS and RFS, of which five were significant at p<0.05 in both OS and RFS (rs10499698, rs10500974, rs13059313,

rs5765061, and rs7312068). Possibly, this suggests that RFS and OS function under different genetic influences. However, genome-wide significance was not achieved in any of the Cox proportional hazards survival analyses, suggesting either a lack of statistical power or a minimal germline genetic component to CRC prognosis at the SNP level. The top SNP associations for OS and RFS are further compared to each other, and to the results of other models used in this chapter, in Section 6.9.

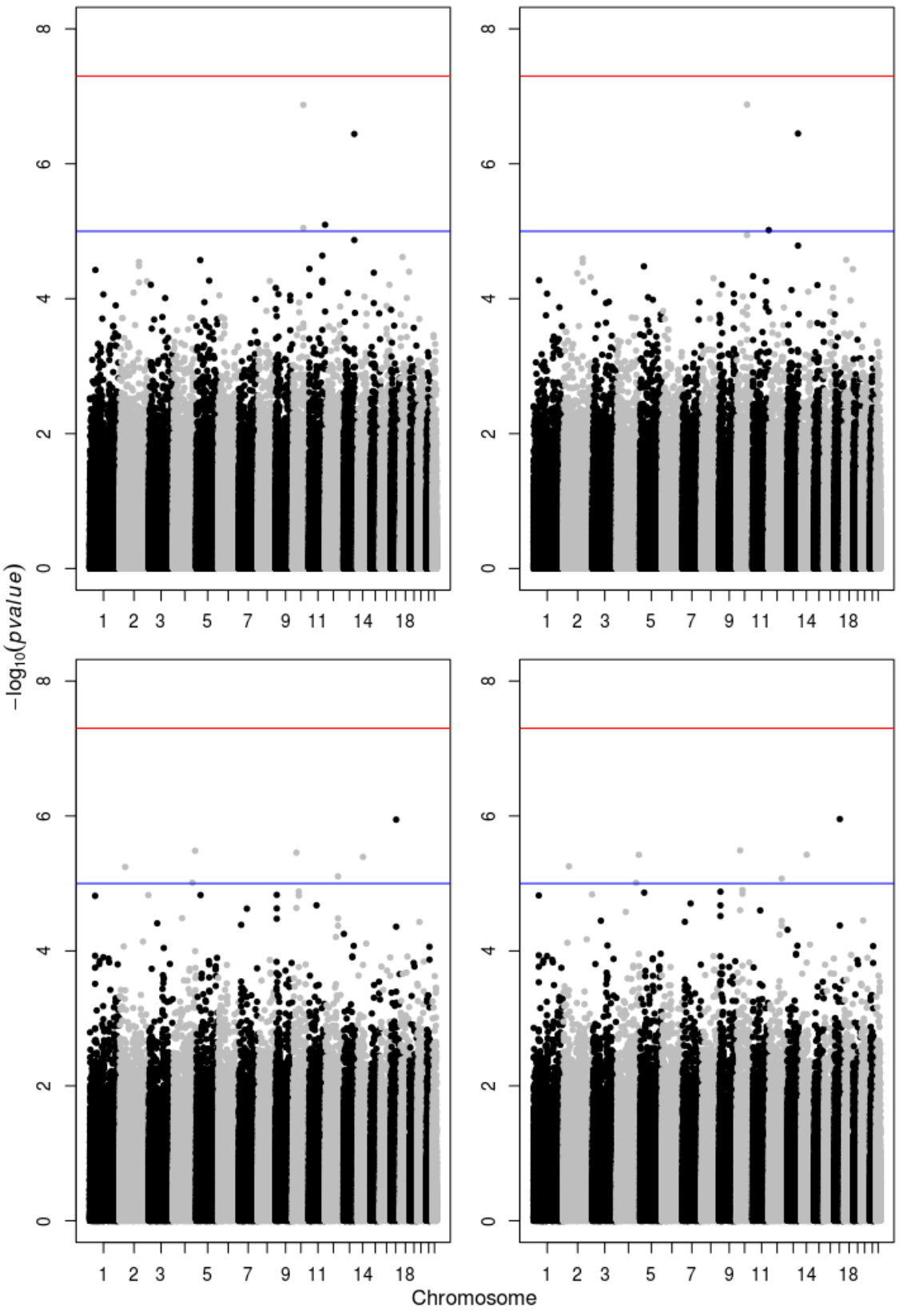


Figure 6.6: Manhattan plots for GWAS of prognosis in the VICTOR trial. Top row is OS, bottom row is RFS. Left column includes blood collection time lag as a covariate, right column does not.

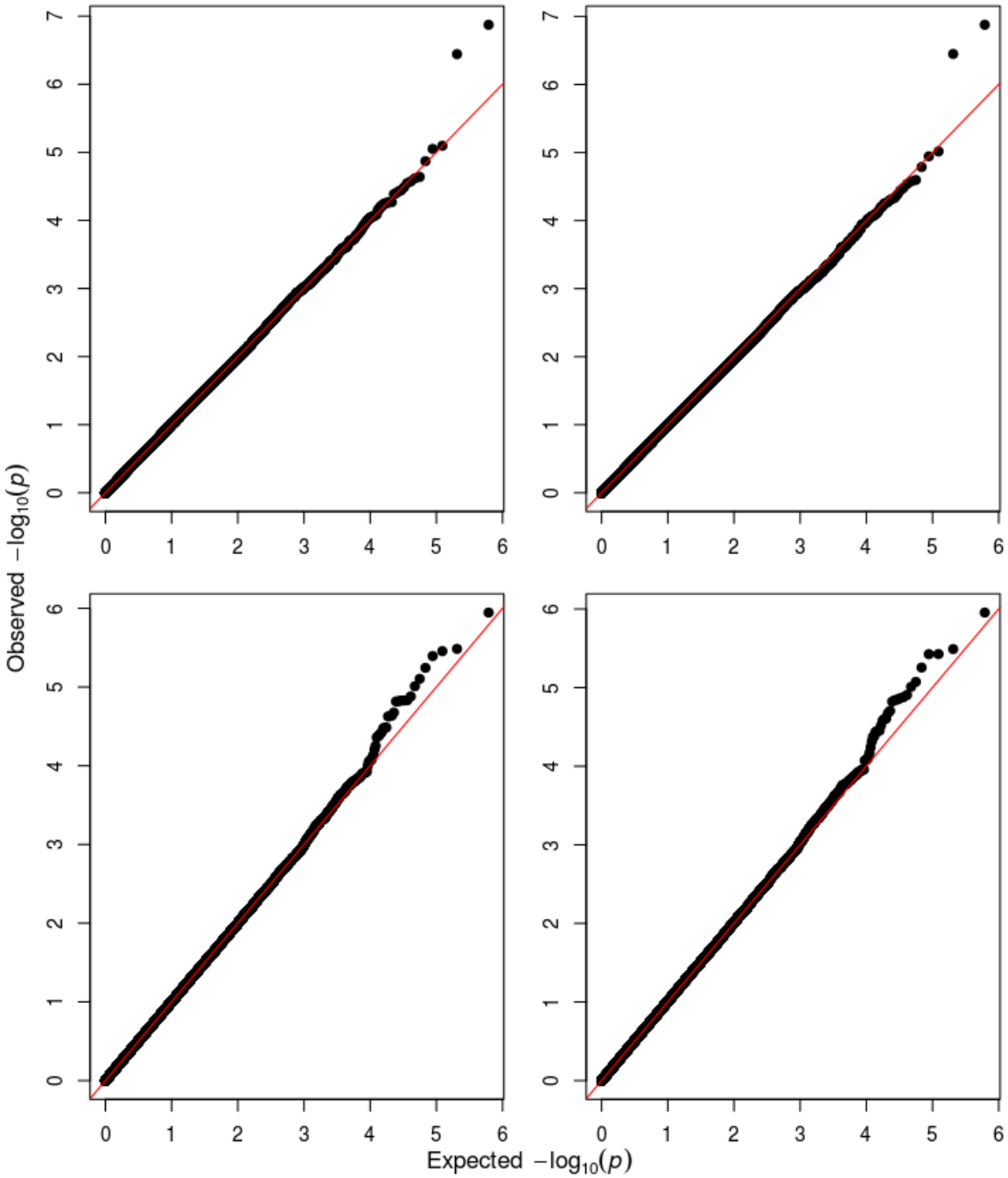


Figure 6.7: QQ plots for GWAS of prognosis in the VICTOR trial. Top row is OS, bottom row is RFS. Left column includes blood collection time lag as a covariate, right column does not. Clockwise from top left, lambda values for inflation were 1.00, 0.98, 1.01, and 1.00

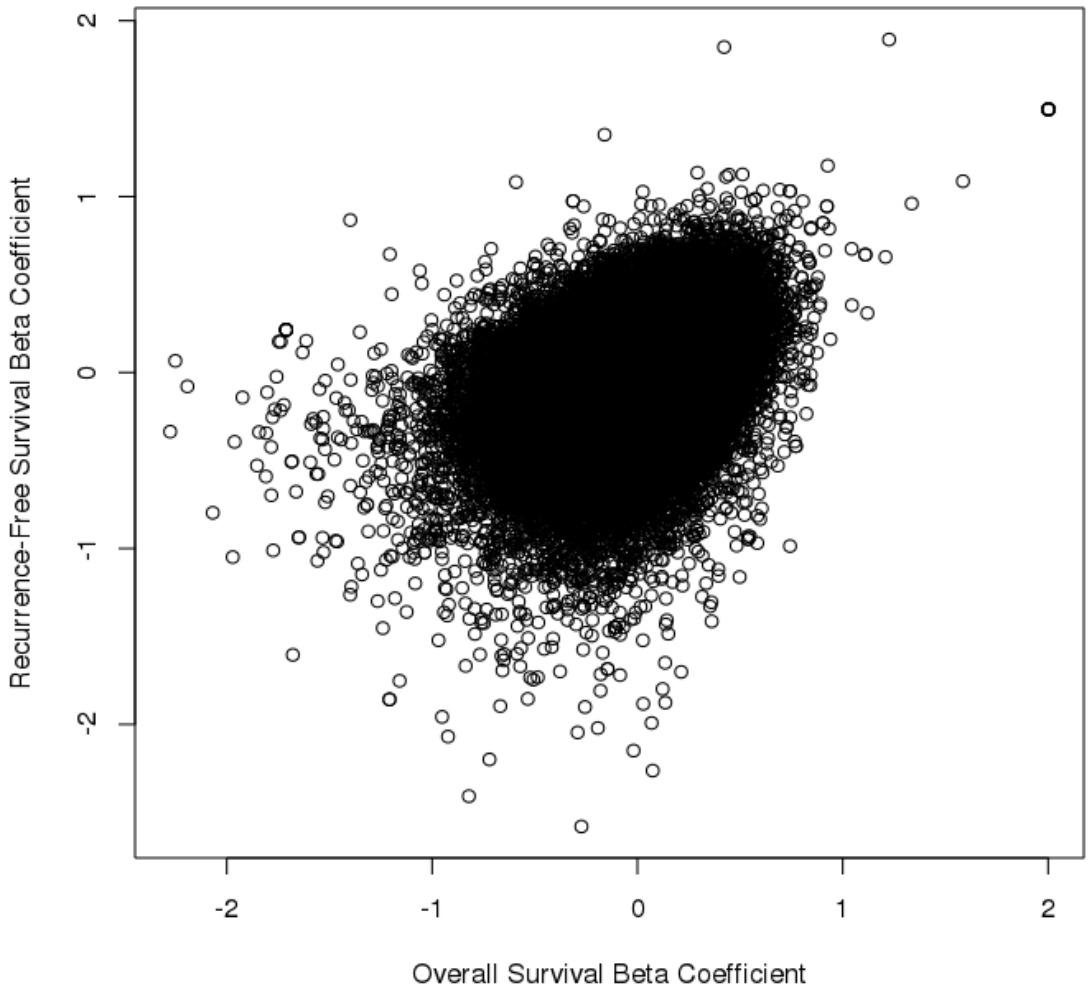


Figure 6.8: The beta coefficients for the OS and RFS Cox proportional hazards models (no lag included as a covariate) show only some correlation (r^2=0.15).

6.7 Prognosis Based on Blood Collection Lag

As mentioned in Section 6.4.4 on covariate selection, the time from surgery to patient blood collection (ie, "lag period") is significantly associated with overall survival when adjusting for clinical covariates in the Cox proportional hazards model. Further, when not conditioning on this lag period, the variable becomes highly significant and has some evidence of violating the proportionality assumption upon which the model is based, whereas other covariates do not. Thus it is necessary to condition on lag period when using the Cox model. However, doing so means that the allele frequency of events (eg, deaths) at a given time from surgery are only compared to that of other patients who are at risk (ie, have had blood collected) at this amount of time after surgery. Given that the lag period is often equal to or greater than the at risk period (ie, lag period at least half of the survival period), this conditioning on survival might lead to a reduction of statistical power. Thus, testing the blood lag period directly as a ("minimum survival") surrogate marker of prognosis might provide insights lost in the Cox proportional hazards model. The adjusted Cox model and the minimum survival analyses are complements, given that the former adjusts for (ie, removes) the effect of lag period on a SNP while testing for the effect of the SNP on survival, whereas the latter tests (ie, isolates) the effect of a SNP on lag period.

Use of blood lag as a surrogate survival marker is first applied to the VICTOR trial. This method is then applied to an entirely separate set of colorectal cancer patients, CORGI, in which lag times are far longer than VICTOR because these patients are recruited as part of a long term consortium whose aim is not to test survival as part of a clinical trial. The most significant associations for each dataset are briefly discussed here, and in Section 6.9 these results are considered as validation for the direction of effect of the lag-adjusted Cox proportional hazards GWAS results from Section 6.6.

6.7.1 Prognosis Based on Blood Collection Lag: GWAS Within VICTOR

Applying this approach to VICTOR, the 918 patients for whom blood was collected were grouped into four quartiles based on the duration of each patient's period from surgery to blood collection. The allele frequency of each SNP was then compared between the first and fourth quartile of patients. Namely, in a normally distributed distribution of lag periods (slightly skewed right) ranging from 7 months to 6 years, with a median lag period of 2 years 4 months and a mean lag period of 2 years 6 months, the first quartile of patients ranged from a 276- to 700-day lag period (0.75 to 1.9 years), while the fourth quartile of patients ranged from a 1,188- to 2,198-day lag period (3.3 to 6.0 years). While of course this dataset excludes patients who died very soon after their surgery and before enrollment in the VICTOR trial, the first and fourth quartiles do represent a 16-month difference in patients' minimum length of survival post-surgery. For a disease for which most disease-related death occurs in the first 3 years, this period is substantial. However, by selecting the extremes, the analysis does face limited power due to reduced sample size.

To assess the effect of genotype on blood lag period (ie, minimum survival), counts were generated for the number of major and minor alleles in each one-quartile patient subset, for each of 307,513 autosomal SNPs. (The SNP subset was slightly different to the Cox analysis due to different SNP exclusions that occurred when shifting the patient inclusion list.) These counts were considered as values in a 2x2 table and analysed under an allelic test for association, applying a Pearson's chi-squared test.

Figure 6.9 shows the results of this analysis. Four associations reach the p<1e-05 suggestive threshold, rs4327776 on chromosome 7 (OR=1.90, p=4.3e-06), rs4794738 on chromosome 17 (OR=0.35, p=1.4e-06), rs2834772 on chromosome 21 (OR=0.28, p=9.7e-06), and rs7675 on chromosome 22 (OR=0.50, p=7.6e-06). The first three of these SNPs are intronic (*DPP6*, *MSI2*, *RUNX1*), while the fourth is in a 3'-UTR (*MI-*

CAL3). Of these genes, *MSI2* and *RUNX1* have a plausible narrative for an impact on CRC survival, given they both have roles in gene regulation. *RUNX1* in particular has a long established role in leukaemia, related to its function in haematopoiesis, and more recently has an identified role in epithelial tissue changes, namely squamous cell carcinomas in mice and breast cancer in humans (191). Further, *RUNX1* has more recently been identified as a tumour suppressor gene in the mouse gastrointestinal track (64). The rs2834772 SNP also has decent counts in each of the tested groups, with 15 minor of 460 total alleles in the short lag group and 49 minor of 458 total alleles in the long lag group. This SNP, however, only achieves modest association and is intergenic, not intronic, when viewed from most protein accessions. The association for this SNP did not replicate when tested in an outside set of CRC patients from the Dunlop group.

On the whole, no convincing associations were found between SNP genotype and blood collection lag in VICTOR, and it is likely that "minimum survival" is not sufficiently powered for the identification of genome-wide significant results, at least when using only 500 patients. SNPs that greatly impact survival, however, should likewise impact minimum survival; thus, in Section 6.9, results from the analysis of blood collection lag in VICTOR are compared to the top results from the Cox proportional hazards model. Further, as the lag period is incorporated into the "full" model in Section 6.10, these results help isolate the impact of different components of the full model, discussed further in that section.

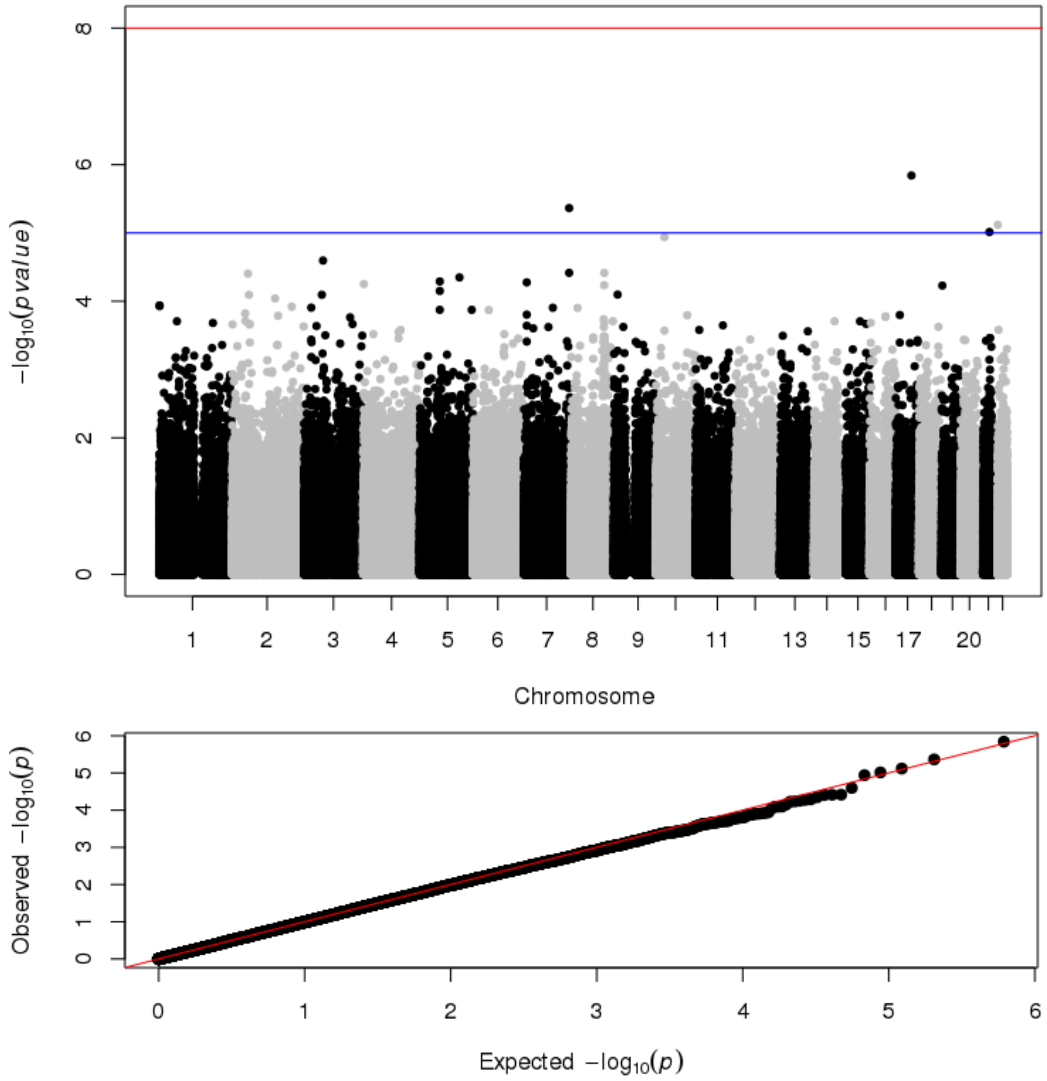


Figure 6.9: GWAS results for the effect of genotype on blood collection lag period (ie, minimum survival) in VICTOR. Results compare VICTOR patients with the first and last quartile of period length from surgery to blood collection, under an allelic model. Lag period is later incorporated into the "full" model in Section 6.10. Lambda of 0.99 suggests no evidence of results inflation.

6.7.2 Prognosis Based on Blood Collection Lag: GWAS Within CORGI

The CORGI (Colorectal Tumour Gene Identification) consortium aims to identify CRC predisposition alleles through multi-centre recruitment of CRC cases and controls UK-wide. Like VICTOR, patients in CORGI often provide blood long after their diagnosis, though this delayed entry is well established, given that recruitment is sometimes retrospective because the study outcome is case/control status, not survival. Thus as a dataset, CORGI provides the advantage of being a long-running study of patients with great differences in minimal survival periods (ie, blood collection lag). Analysis of minimum survival in CORGI might provide markers of genome-wide significant effect or act as validation of markers identified in VICTOR by the Cox method.

As part of previous work in the Tomlinson lab, genotype data for 555,352 SNPs on the Illumina HumanHap550 BeadChip Array were available for 620 CORGI CRC cases, for whom date of diagnosis and blood arrival data were immediately available for 358. Diagnosis date or blood arrival date (whose difference was defined as the lag period) were estimated manually for a further 127 patients by looking at other clinical variables such as date of birth together with age at diagnosis. This in total formed a 485-patient cohort for analysis. The blood collection lag period for this cohort was severely skewed to the right, with a lag period range of 0 months to 58 years, mean of 9 years 4 months and median of 7 years 9 months. As for VICTOR, an allelic comparison of patients in the first lag-length quartile (0.0 to 4.1 years; n=121) and fourth lag-length quartile (11.9 to 58 years; n=121) was made.

Results of the analysis were not inflated (lambda=1.01). The top hit in this analysis was rs6731901 on chromosome 2 (OR=0.40, p=1.5e-06), which is intronic in *MYO1B*, a myosin isoform that may play a role in cell motility (124) and whose expression has been shown to be increased in acute myeloid leukemia (212). Unfor-

tunately, while the SNP is not genotyped among VICTOR patients, a perfect proxy (rs11892998) shows no significance in the adjusted Cox proportional hazards model. Other associations at p<1e-05 include rs2017653 (intergenic, OR=0.30, p=4.1e-06), rs2280851 (*HHLA1* splice site missense, OR=0.36, p=8.1e-06) and rs13036466 (29kb upstream of *GATA5*, OR=0.40, p=8.4e-06). Although this locus is intriguing, the *HHLA1* SNP was also not significant in the VICTOR Cox model. No results reached genome-wide significance, at least partly a reflection of statistical power. A Manhattan plot of the results is not shown because the results are better interpreted as a validation set for the direction of effect of the top associations in the Cox proportional hazards model, as are compared in Section 6.9, and because these CORGI data are not incorporated into the "full" model in Section 6.10.

6.8 Prognosis Based on VICTOR/QUASAR2 MAF Differences

The QUASAR2 trial provides a strong dataset for CRC survival analysis that will eventually be analysed, but whose survival period does not mature until 2014. However, this dataset can still contribute to a survival analysis when we recognise that, contrary to VICTOR, bloods were collected from patients in QUASAR2 without any significant lag period (ie, no minimum survival period). In a sense, then, we can treat all QUASAR2 patients as cases in terms of their minimum survival period and all VICTOR patients from whom blood was collected (ranging from a 7 month to 6 year minimum survival period) as controls; this also provides insights into the "full" model in Section 6.10. Further, as the ethnic origins of these populations overlap strongly and both are from trials of stage II/III CRC patients, it would seem unlikely that population structure should hinder an analysis involving both patient sets. One confounder might be genotyping differences given that different array platforms were

used; however, as noted in regards to the genome-wide imputation of QUASAR2, the pooling of data from the Hap300 array (used in VICTOR) and other QUASAR2 arrays did not lead to inflation of results. Thus, in this section, the allele frequencies of the intersect of genotyped variants in the VICTOR and QUASAR2 trial are compared between the two trials, in search of variants with modest significance in the Cox proportional hazards model and others.

For the analysis, the number of minor and major alleles for each of 305,974 autosomal SNPs genotyped in both QUASAR2 and VICTOR (and individually passing quality control) were analysed. First, allele counts were generated separately genome-wide for the 955 included QUASAR2 patients and the 918 included VICTOR patients. Next, at each SNP, a 2x2 table was formed of the number of major and number of minor alleles among QUASAR2 (cases) and VICTOR (controls) patients. Then, an odds ratio was calculated for the allelic imbalance of each SNP and a Pearson's chi-squared significance test was performed. In essence, an odds ratio greater than 1 suggests that an allele reduces survival following CRC diagnosis, as QUASAR2 patients (with shorter minimum survival times) are more likely to hold this allele than VICTOR patients from whom blood was collected (with longer minimum survival times).

Figure 6.10 shows the results of this analysis. One SNP (rs4862110) was removed from the analysis as it showed a p-value of 1.1e-17 but has been previously noted to genotype poorly on the Hap610 array (184). Seven SNPs fell below the p=1e-05 threshold (rs12406238, rs2319335, rs632916, rs11110069, rs2236282, rs9901079, rs3787950), none of which was significant at $p<0.05$ in the Cox proportional hazards model; a literature search of these variants and their genes provided no strong rationale for CRC survival, and due to their modest significance, no additional follow-up was performed. As with the models directly testing blood lag (Section 6.7), the VICTOR vs QUASAR2 MAF comparison is best used as validation of the direction of

results from the Cox model (Section 6.9) and as a means of isolating components of the "full" model (Section 6.10).

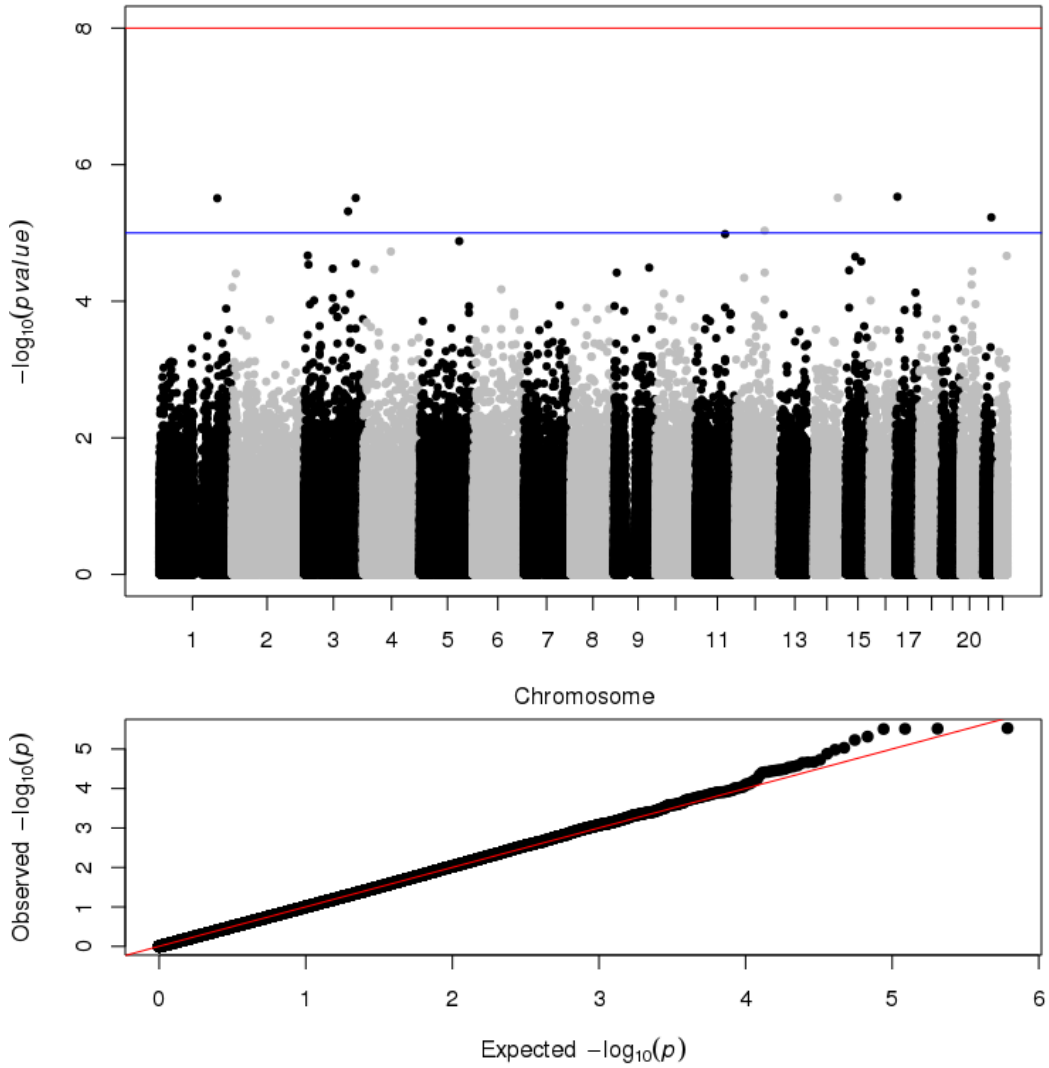


Figure 6.10: GWAS results for the difference in MAF between VICTOR and QUASAR2, a potential approximation for a survival analysis, given that VICTOR patients act as controls due to a 7 month to 6 year blood collection lag (ie, minimum survival period) and QUASAR2 patients act as cases because their bloods were collected without delay. Lambda of 1.01 suggests no evidence of results inflation.

6.9 Comparison of Results Across Methods

In addition to our gold standard Cox proportional hazards OS and RFS models, the GWAS based on CORGI blood collection lag period provides an independent surrogate measure of survival, and the GWAS based on VICTOR blood collection lag period and the genome-wide comparison of VICTOR/QUASAR2 allele frequency differences provide partially independent measures. While most of these models are likely underpowered alone for the detection of alleles that increase or decrease patient prognosis when applying a genome-wide significance threshold, those SNPs that consistently increase or decrease survival across models form a shortlist of top hits that should be prioritised for future validation.

Table 6.3 lists 44 SNPs that showed a consistent direction of effect (irrespective of p-value) across models. To identify these variants, all SNPs (n=609) with a p-value below 0.001 in either the VICTOR OS or RFS Cox proportional hazards model (our gold standard analyses) were extracted, then four sets of inclusion criteria were applied to determine consistency across models. For the first two sets, all SNPs with a *modest* effect size in the same direction – with OR/HR>1.1 or its inverse, with OR/HR<0.91 – in either the OS or RFS Cox model and in all three of the other models were extracted (n=6). For the second two sets, all SNPs with a *large* effect size in the same direction – with OR/HR>1.5 or its inverse, OR/HR<0.67 – in either the OS or RFS Cox model and in at least one of the other models were extracted (n=37). A few of these SNPs seem of particular interest, based on gene function: *PLK2* plays a role in normal cell division, *ADAMTSL1* plays a role in the extracellular matrix (possibly affecting metastasis), *CDH8*, *CDH11*, *NRXN1*, *PCDH7*, and *PCDH17* mediate cell adhesion (again possibly affecting metastasis), *SMARCA2* is a regulator of transcriptional activity, *DKK1* inhibits the WNT signaling pathway, *MBL2* activates the innate immune system, and *NAP1L1* regulates cell proliferation.

SNP	Chr	Coordinate	Ref	Eff	Gene(s)	Function	COXPH OS P-value	COXPH OS HR	COXPH RFS P-value	COXPH RFS HR	VICTOR Lag Period P-value	VICTOR Lag Period OR	CORGI Lag Period P-value	CORGI Lag Period OR	QUASAR2/VICTOR MAF P-value	QUASAR2/VICTOR MAF OR
OR>1.1 in OS and/or RFS and OR>1.1 in all other models																
rs1052607	1	46272113	G	A	MAST2	intronic	3.8E-05	2.3	1.3E-01	1.5	3.8E-01	1.3	7.1E-01	1.2	4.6E-01	1.1
rs17315098	5	56938838	A	G	ACTBL2; PLK2	intergenic	7.7E-04	1.8	6.8E-01	1.1	2.6E-01	1.2	6.6E-01	1.1	1.8E-01	1.1
rs10810875	9	17912697	G	A	SH3GL2; ADAMTSL1	intergenic	2.1E-01	1.3	4.5E-04	2.0	1.1E-01	1.4	7.5E-01	1.1	3.5E-01	1.1
rs8045301	16	63149304	G	A	CDH8; CDH11	intergenic	9.8E-04	1.6	6.9E-01	1.1	2.6E-01	1.2	4.1E-01	1.2	2.1E-01	1.1
rs7208097	17	2891528	A	G	RAP1GAP2	intronic	1.5E-04	1.6	6.5E-01	1.1	1.1E-01	1.3	3.8E-01	1.2	2.3E-01	1.1
OR<0.91 in OS and/or RFS and OR<0.91 in all other models																
rs3794042	11	60453460	G	A	LINC00301	ncRNA intronic	1.0E-03	0.6	2.7E-02	0.7	3.2E-01	0.9	3.6E-01	0.8	1.6E-01	0.9
OR>1.5 in OS and/or RFS and OR>1.5 in at least one other model																
rs12404118	1	66345841	A	G	PDE4B	intronic	7.0E-04	2.0	2.3E-02	1.7	5.7E-01	1.1	7.2E-02	1.8	3.7E-01	0.9
rs7549052	1	214016229	A	G	RPS6KC1; LINC00538	intergenic	5.9E-01	1.1	5.0E-04	1.9	9.4E-01	1.0	1.0E-01	1.6	5.1E-01	0.9
rs1454312	2	33109488	G	A	LINC00486	ncRNA intronic	1.8E-01	1.3	8.6E-05	2.3	2.9E-01	0.7	2.3E-01	1.5	9.0E-02	0.8
rs12615200	2	45298018	G	A	SIX2; UNQ6975	intergenic	1.9E-01	1.3	3.8E-04	2.0	9.5E-01	1.0	1.5E-01	1.6	4.0E-01	0.9
rs883834	2	45310615	A	G	SIX2; UNQ6975	intergenic	4.9E-03	1.8	5.7E-06	2.5	8.9E-01	1.0	2.7E-01	1.5	8.7E-01	1.0
rs2049735	2	51605593	G	A	NRXN1	intergenic	2.6E-01	1.2	7.5E-04	1.7	1.6E-01	1.3	7.2E-02	1.5	7.0E-01	1.0
rs17728568	2	115203927	G	A	DPP10	intronic	8.1E-05	2.9	3.4E-01	1.4	1.4E-01	1.7	2.4E-01	1.8	1.5E-03	0.6
rs6433636	2	177599390	A	G	LOC375295; HNRNPA3	intergenic	1.1E-01	1.4	8.9E-04	1.9	5.0E-02	1.6	3.0E-02	2.0	9.9E-01	1.0
rs977103	3	67456931	A	G	SUCLG2	intronic	2.6E-01	1.2	3.9E-05	2.1	7.6E-02	0.7	1.6E-01	1.5	8.1E-01	1.0
rs1435796	4	29918340	A	C	MIR4275; PCDH7	intergenic	5.8E-01	0.8	8.0E-04	2.4	7.5E-01	1.1	1.3E-01	1.8	8.5E-01	1.0
rs7706821	5	67274234	G	A	CD180; PIK3R1	intergenic	1.1E-04	1.7	3.8E-02	1.4	4.4E-01	1.1	9.4E-03	1.7	8.9E-01	1.0
rs12661179	6	10471014	A	G	MIR5689; GCNT2	intergenic	8.9E-05	3.5	1.3E-01	1.9	7.9E-01	0.9	3.6E-01	1.8	7.9E-04	0.5
rs6905047	6	76554829	A	G	MYO6	intronic	1.6E-03	1.7	4.7E-04	1.9	2.6E-01	1.2	1.2E-01	1.6	7.4E-01	1.0
rs7805713	7	66893183	A	G	STAG3L4	intergenic	4.4E-01	1.3	4.8E-04	2.6	6.1E-01	0.8	1.7E-01	1.9	8.6E-01	1.0
rs4734889	8	106761674	G	A	ZFPM2	intronic	1.4E-01	1.3	5.6E-04	1.8	9.7E-01	1.0	9.3E-02	1.5	1.4E-01	1.1
rs12345035	9	1212087	G	A	DMRT2; SMARCA2	intergenic	1.4E-04	2.3	9.7E-01	1.0	9.1E-01	1.0	1.4E-01	1.6	9.0E-01	1.0
rs10867816	9	83619861	A	G	TLE4; TLE1	intergenic	8.0E-04	2.0	1.4E-01	1.4	9.6E-02	1.6	3.0E-01	0.7	3.7E-01	0.9
rs11816856	10	54168396	G	A	DKK1; MBL2	intergenic	7.6E-04	2.2	1.8E-01	1.5	1.8E-01	1.5	1.3E-01	1.8	3.9E-01	0.9
rs12267240	10	88229480	G	A	WAPAL	intronic	4.9E-04	1.9	6.2E-03	1.8	6.3E-02	1.6	--	--	5.5E-01	1.1
rs7477274	10	98539567	G	A	PIK3AP1; LCOR	intergenic	7.4E-02	1.5	6.0E-04	2.0	3.9E-01	0.8	1.3E-02	2.1	7.7E-01	1.0
rs711129	12	76568088	G	A	NAP1L1; BBS10	intergenic	4.4E-04	2.3	8.5E-01	0.9	9.6E-02	1.7	1.1E-01	1.8	5.1E-01	0.9
rs10774518	12	118775846	A	C	TAOK3	intronic	2.4E-04	1.7	1.7E-01	1.3	9.6E-01	1.0	7.0E-02	1.5	6.8E-01	1.0
rs10219804	13	58275814	A	C	PCDH17	intronic	5.3E-01	1.2	7.4E-04	2.5	3.8E-01	1.4	4.3E-01	1.5	2.5E-01	0.8
rs2402172	14	92120323	C	A	CATSPERB	intronic	1.6E-01	1.3	7.8E-05	1.9	4.7E-01	1.1	4.5E-02	1.6	9.2E-01	1.0
rs6496026	15	92973387	A	C	ST8SIA2	intronic	1.7E-04	2.1	3.6E-03	1.9	4.4E-01	0.8	1.8E-01	1.7	9.8E-01	1.0
rs9933802	16	53986635	C	A	FTO	intronic	3.4E-04	2.0	7.0E-01	1.1	5.0E-01	0.9	1.6E-01	1.5	7.6E-02	1.2
rs11640811	16	54000381	A	G	FTO	intronic	2.8E-04	2.2	4.0E-01	1.3	9.0E-01	1.0	3.1E-01	1.4	1.9E-03	1.5
rs2576542	16	54037453	A	G	FTO	intronic	8.7E-04	1.9	5.9E-01	1.1	3.3E-01	0.8	4.0E-02	1.8	8.4E-01	1.0
rs17241017	16	54043258	A	G	FTO	intronic	1.3E-04	2.4	2.8E-01	1.4	6.6E-01	0.9	5.4E-02	2.0	3.2E-01	1.2
rs7221480	17	73370128	C	A	GRB2	intronic	4.6E-01	1.3	2.2E-04	2.8	5.9E-02	2.3	2.1E-01	0.5	2.6E-01	0.8
rs10402666	19	48560178	A	G	PLA2G4C	intronic	9.8E-04	2.1	7.4E-01	0.9	8.1E-02	1.6	5.7E-01	0.8	2.6E-01	1.2
rs2050406	20	16563735	A	G	KIF16B; SNRPB2	intergenic	9.2E-04	1.9	2.5E-01	1.3	3.8E-01	0.8	5.0E-02	1.7	1.2E-01	0.8
rs6021441	20	49844950	A	G	KCNG1; NFATC2	intergenic	4.3E-01	1.1	6.2E-04	1.6	4.8E-01	1.1	1.3E-02	1.6	3.5E-01	0.9
rs9305294	21	27147453	G	A	GABPA; APP	intergenic	8.1E-02	1.5	4.6E-04	2.2	6.0E-02	1.7	7.0E-01	0.9	7.3E-01	1.0
rs11909037	21	27148423	G	A	GABPA; APP	intergenic	8.1E-02	1.5	4.6E-04	2.2	6.0E-02	1.7	7.0E-01	0.9	7.5E-01	1.0
OR<0.67 in OS and/or RFS and OR<0.67 in at least one other model																
rs7782262	7	90509611	C	A	CDK14	intronic	7.6E-04	0.5	3.4E-02	0.6	8.2E-01	1.0	6.8E-02	0.7	2.9E-01	0.9
rs1982221	12	95256436	G	A	KRT19P2; NDUFA12	intergenic	4.0E-01	0.9	4.6E-04	0.5	8.8E-02	1.3	3.1E-02	0.7	9.8E-01	1.0

Table 6.3: Among the 609 SNPs with p<0.001 in either the OS or RFS Cox proportional hazards model, 43 SNPs showed consistent direction of effect across the various prognosis models described in this chapter. Four sets of criteria were applied to identify these SNPs, as described in text. All reported SNPs had HWE p-values above 1e-04 in all datasets. Beta values coloured pink are those that achieve the threshold of the set, for the two sets in which a SNP only needed to achieve the effect size threshold in at least one model in addition to a Cox model.

6.10 Novel Method for Survival Analyses of Truncated Datasets

The models described in this chapter each make use of factors not utilised in the Cox proportional hazards model: measuring collection time lag instead of survival time might increase statistical power by including more patients in allele frequency comparisons (Section 6.7) and a VICTOR/QUASAR2 per-SNP MAF comparison might identify alleles that were fully or partially removed from VICTOR due to the blood collection lag (Section 6.8). These models are crude – by incorporating only one factor they are prone to identifying confounders (eg, population differences) as survival associations. A more robust model would maintain the hazard function that predates even the Cox proportional hazards method, while also incorporating blood lag period, VICTOR/QUASAR2 MAF differences and other factors particular to the truncated VICTOR dataset. In order to build such a model, I worked with a Department of Statistics DPhil student, Tristan Gray-Davies. We together designed a novel survival model that we hoped would be more robust for the analysis of truncated survival data, Tristan coded the model and ran it using Statistics with R and I plotted and analysed the output. This section describes this model, designated the "full" model because it fully models the hazard function instead of relying on a partial likelihood-based model such as Cox proportional hazards and because it incorporates additional factors particular to the VICTOR dataset.

The goal of any genetic survival analysis is to identify the odds of survival with genotype "G" versus the odds of survival with genotype "g". Although the period of survival is censored at the end of follow-up, the survival rate is continuous; however, for ease of understanding, one can consider survival in terms of one timepoint "T", for which the odds of surviving to T is compared for patients with genotype G versus genotype g – and then the analysis is repeated for all timepoints T. As described

previously, in the standard Cox proportional hazards model, a partial likelihood is formed in which the hazard function is never specified, but rather drops out when an odds ratio is formed between a numerator specifying the null model and a denominator specifying the null hazard function plus covariates. With genotype as a covariate, G is compared to g in an odds ratio, adjusted for other covariates.

For the "full" model, we identified three sets of information that could not be incorporated into the Cox proportional hazards model but that should be informative in understanding how the odds of survival are affected by genotype G vs g, thus potentially improving statistical power for detecting an association if included in the model:

1. Allow for an inconsistent effect of covariates over time (which breaks the proportional hazards assumption), specifically for blood collection lag period, which showed some evidence of an increasingly negative effect (beta) over time by Schoenfeld residuals and that was a nearly significant concave down curve when not conditioning on lag period (Section 6.4.4).

2. Use clinical data from all 2,383 VICTOR patients when identifying the effect (ie, beta coefficient) of covariates (eg, disease stage, but also blood collection lag period), instead of just those 918 VICTOR patients with both full clinical and genetic data who were included in the Cox proportional hazards model, allowing for a more precise estimate of the effect of included variables.

3. Incorporate allele frequencies from QUASAR2 (a similar patient set, but which had no lag in blood collection) as a comparison group, in order to better detect alleles of true effect that were fully or partially lost due to the blood collection lag period – an advantage that could not be achieved through use of lag period as a covariate.

In order to incorporate this information, we made the following substantial changes

to the Cox proportional hazards model (some of which are a reversion to models used prior to the invention of proportional hazards):

1. Specify the baseline hazard (as opposed to allowing it to cancel by proportional hazards), approximated according to the Breslow method for estimating the average (cumulative) hazard (26); this allows the lag period to be incorporated into the hazard function as a time-dependent variable.

2. Use clinical data from all 2,383 VICTOR patients when identifying the beta coefficients specified in the Breslow estimator.

3. Use QUASAR2 genotype data to specify the expected distribution of alleles for each SNP at the time of surgery.

In essence, the "full" model is an amalgamation of the analyses shown in the Manhattan plots in Figure 6.6 for the OS Cox model, Figure 6.9 for lag period and Figure 6.10 for the VICTOR/QUASAR2 MAF comparison. The results of this full model are shown in Figure 6.11. Only overall survival was analysed, adjusted for age, stage and gender and with blood collection lag period incorporated within the average hazard. Recurrence-free survival, alternatively, could not be analysed with the model because the full model could not allow for patients who experienced an event (ie, recurrence) prior to becoming at risk for a genetic analysis (ie, prior to blood collection) – a peculiarity specific to the VICTOR dataset. Although two SNPs reach genome-wide significant or suggestive thresholds – rs9514816 on chromosome 13 (4.68e-08) and rs4934251 on chromosome 10 (1.91e-07), both of which reached suggestive significance in the Cox proportional hazards model – significant inflation was observed across the results, discrediting the believability of these significance levels. Neither of these two variants demonstrated a consistent effect across models (Section 6.9).

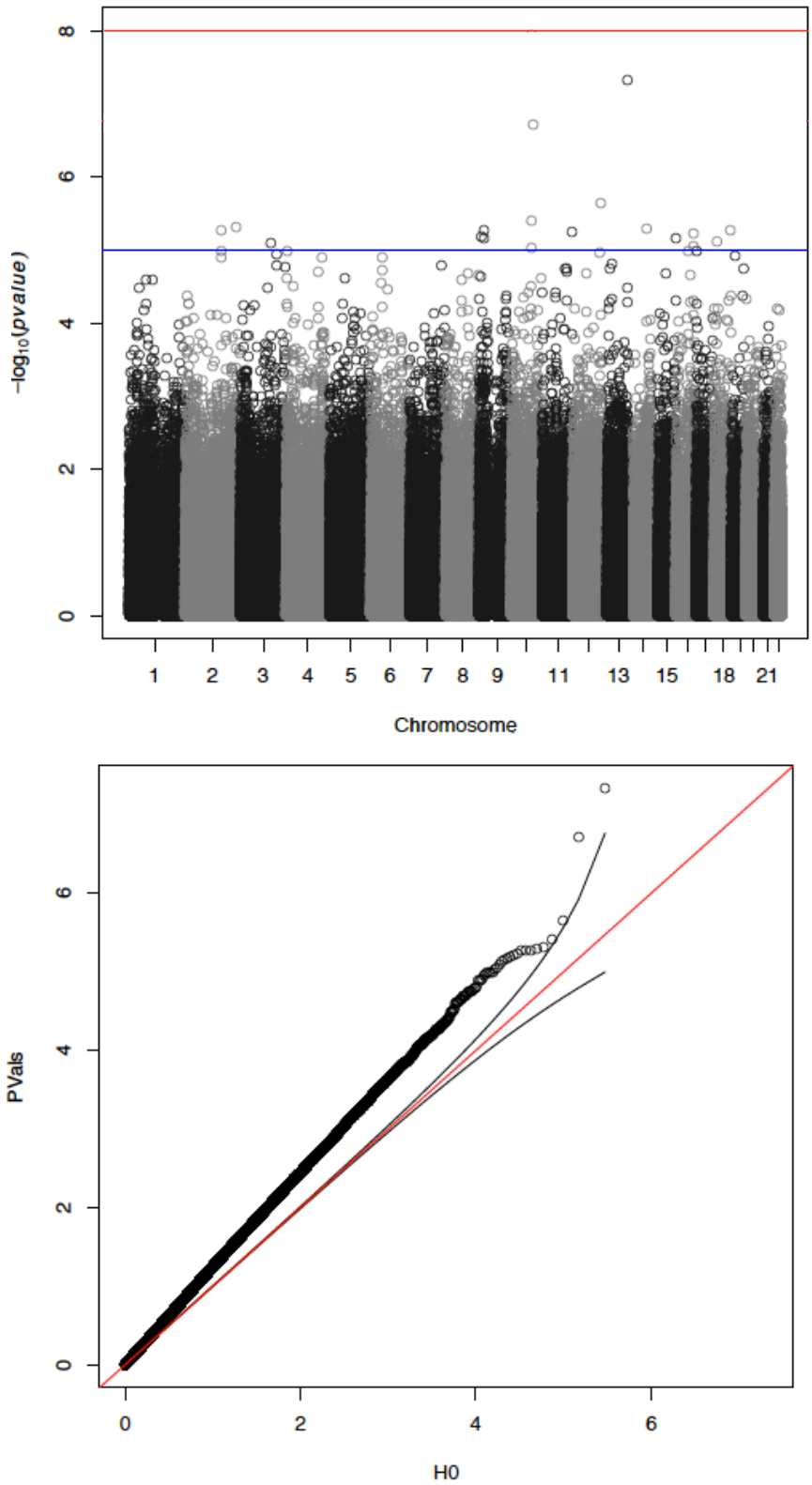


Figure 6.11: GWAS results for overall survival in VICTOR using the "full" model to specify a baseline hazard, incorporate clinical data from all 2,383 VICTOR patients and apply genotype data from the QUASAR2 trial. Lambda of 1.31 suggests that significance values for associations are inflated.

Intriguingly, while incorporating much additional information, the full model did not equally inflate all associations. Possibly, this suggests that the model is not fundamentally flawed, but rather that the extra information it includes does in fact increase statistical power, while some particular aspect of the model causes inflation. If this were the case, then the full model could be used to better identify true associations among true null associations, given that the flawed component were fixed. In discussions with Tristan, one likely cause of inflation is that the SNP allele distributions from QUASAR2 are too strongly imposed, causing the model to attribute all allele frequency differences between the VICTOR and QUASAR2 patient sets as related to survival/lag; as seen in Section 6.8, multiple MAF differences on the order of 5e-06 p-value exist between the datasets, and these p-values are normally distributed (meaning many random associations of modest significance exist). As the model currently assumes that the QUASAR2 allele frequencies are correct, and that VICTOR allele frequency differences are entirely the result of lag and survival periods, the significance of almost all SNPs would be inflated. If the model instead assumed that the QUASAR2 and VICTOR allele frequency distributions were two samples from a larger UK CRC patient population, then only large MAF differences should substantially alter significance. Re-coding the model for further analysis is the subject of current work.

6.11 Discussion

In this chapter, I have outlined the challenges related to a survival analysis of germline markers in the VICTOR trial and performed and compared analyses using four models: the Cox proportional hazards model, an allele frequency comparison of patients with long and short blood lag collection (in both VICTOR and CORGI), an allele frequency comparison of patients in VICTOR and QUASAR2, and a novel "full" model

that provides more information than can be incorporated into the Cox model. While some SNPs achieve p<5e-08 significance in single models – such as rs6731901 in the CORGI allele frequency comparison by collection lag time – none of the highly significant SNPs in the non-standard models achieve modest significance in the Cox model, our gold standard.

Possibly, individual SNPs do not impact patient prognosis significantly enough so as to be detected by a 1,000 patient GWAS. However, the severe left truncation observed in the VICTOR trial – the result of a long lag in blood collection time – suggests that this analysis was less powered than most studies of similar size. As such, beyond the standard Cox model, I tried to utilise this deficit in multiple ways. First was to use the lag period to stratify patients into groups with long and short "minimum" survival periods. Next was to use the MAF observed in QUASAR2, a similar patient set, as our expected allele frequency and to identify SNPs with significantly different allele frequencies in VICTOR. Third was a less crude approach to incorporate these factors, designated the "full" model. This model incorporated covariate data from all 2,383 VICTOR patients, the allele frequency of QUASAR2 and the lag period of for blood collection, but inflated the significance of associations.

Some previous work has aimed to identify genetic associations in light of left truncation, namely Bolton et al, who dealt with a high case prevalence resulting from recruitment at variable time periods following ovarian cancer diagnosis by measuring time from diagnosis that was adjusted for (ie, conditional on) the period from diagnosis to recruitment (21). This form of recruitment-based left truncation exists in VICTOR but is seconded by the lag in blood collection that impacts a genetic study. Further, although the Bolton work was for a continuous period of left truncation, outcome was case/control, not continuous time-to-event periods. Interestingly, though, it has been shown that hazard ratio effect size estimates for survival analyses are not biased by presence of prevalent cases (recruited after diagnosis), so long as

proportional hazards assumptions are not violated (10). This means that effect estimates in VICTOR with sufficient statistical power so as to be significant should be accurate (though power was reduced for all estimates in VICTOR because of patients lost during the lag period). The median lag period from this study was 1 year 7 months, compared with a median 2 years 3 months in VICTOR, so it is reasonable that the result should apply to VICTOR. One concern, however, is that the blood lag period showed some evidence of violating proportional hazards, in which case beta estimates could be either over- or under-estimated, as determined by when the slope of the lag period differs from the slope of the baseline hazard function.

It remains possible that germline markers of CRC prognosis exist, including those within reach of a 1,000-patient GWAS. As reported in Table 6.3, at least 43 SNPs showed consistent evidence for an impact on survival periods, and these should be prioritised for future follow-up work. While further model building related to the VICTOR trial might provide stronger evidence of an association, it is unlikely that it will achieve the power of a non-truncated dataset. Thus, the survival analyses for the QUASAR2 dataset – scheduled for 2014 – are eagerly awaited, both to identify new markers and to compare to the models of this chapter as validation of its methods.

Chapter 7

Discussion

This thesis used the germline to investigate two areas related to colorectal cancer treatment – toxicity to 5-FU/capecitabine and prognosis after diagnosis – with the aims of using genetics to improve personalisation of care and expanding research methods in the field. With approximately 40,000 UK and 140,000 US colorectal cancer diagnoses expected for 2013, selection of appropriate treatment decisions will have a wide impact on treatment decisions and appropriate allocation of healthcare resources.

In the study of capecitabine/5-FU-related toxicity, this thesis identified multiple new markers of toxicity and revised our understanding of previously studied variants. Only 4 of 36 variants previously tested for an association with 5-FU-related toxicity – or 4 of 12 tested variants included in commercial kits for predicting such toxicity – were found to have strong evidence for an effect on toxicity. The remainder did not, despite 16 having been reported at $p<0.10$ in at least one previous publication. Of those with strong evidence, two – *TYMS* 5'VNTR and 3'UTR – may not be functional, as once assumed, but rather tag rs2612091 or other variants tagged by this SNP, possibly at *ENOSF1*. The other two previously tested variants are both rare, functional *DPYD* variants – *2A and 2846 – whose strong toxicity associations in

QUASAR2 and other meta-analysed studies provide a priori evidence for the inclusion of 15 additional putative DPYD Deficiency Syndrome causal variants as a functional class. In addition, three new *DPYD* variants that affect risk of 5-FU toxicity have been discovered – rs12132152, rs7548189 and A551T – of which the last is a putative DPYD Deficiency Syndrome causal variant (205), further supporting the notion of this functional class. The 5 *DYPD* and 1 *ENOSF1 / TYMS* variants together provide an 0.66 area under curve for predicting capecitabine-related toxicity, an improvement on commercially available kits. Additionally, this thesis has provided a summary, based on 16 studies, of the differing incidence of specific toxicity phenotypes and SNP effect sizes across 5-FU-based regimens, noting that associations in one regimen are not clearly generalisable to other regimens, particularly for combination therapy regimens such as FOLFOX.

In the study of prognosis, this thesis identified that the VICTOR dataset, while reduced in statistical power due to the absence of genetic information for some patients experiencing deaths, is unlikely to be inherently biased as a result of its blood collection lag, meaning that the effect size estimates this dataset produces should be accurate but less likely to achieve significance. Among over 300k SNPs tested, neither the gold standard Cox proportional hazards method nor multiple non-standard analyses identified variants with genome-wide significant associations with prognosis, meaning that we can be confident that very strong markers of CRC survival are unlikely to exist at this SNP density. This thesis provided a list of top associations for replication in further similar studies, as this set of variants may contain true markers that require larger sample size to detect.

In addition to genetic biomarker results, this thesis demonstrated the capacity of certain genetic methods. Meta-analysis is a common method used to combine data across studies, particularly in medical sciences, but the method is mostly used to test one outcome or one polymorphism (or set of polymorphisms, such as GWAS), both

of which are pre-defined. This thesis showed that a two-step analysis is possible in which the study protocol is determined after preliminary analysis of collected data, in this case identifying the association testing protocol (which patient sets to combine) based on an analysis of toxicity incidence and per-regimen SNP effect sizes from the collected data. Although next generation sequencing technology has evolved rapidly even in the life of this thesis – more easily allowing for, as example, the use of bar codes to tag individual patients – this report demonstrates that even personal mutations can be identified when performing amplicon sequencing in equimolar patient pools, as has been previously proposed (22). The work described here also corroborates the continued usefulness of SNP array data and GWAS in an era of expanding sequencing use, particularly when sample size is high, outcome data are carefully collected and imputation to high density is performed; previous candidate-based and genome-wide analyses of chemotherapy toxicity have largely achieved lesser levels of significance. Considering rarer variants, this thesis proposes a method for estimating the effect size on toxicity of rare variants, such as those in *DPYD* – specifically, as a functional class. The zCall software was also successfully implemented in calling such rare variants.

The work in this thesis was greatly aided by new methods in the field, in addition to the sequencing and imputation methods mentioned above. Functional annotation tools in particular have improved interpretation of association testing results, now providing nearly high-throughput capacity. Whereas annotation tools for protein change have existed for about five years – based mostly on base pair conservation and predicted protein structure, such as SIFT, PolyPhen and MutationTaster – newer initiatives such as ENCODE and The Cancer Genome Atlas have provided genome-wide experiment-based functional annotation. Although not usually capable of definitive research answers, these tools quickly help prioritise results and guide follow-up analyses, such as for the rs2612091 *TYMS/ENOSF1* SNP described in the 5-FU/capecitabine pathway analysis (Chapter 4). The exact functional mechanism

by which this SNP affects toxicity remains unknown, but publicly available resources helped tip our understanding toward a role involving *ENOSF1*.

Despite ever-improving methods, research remains limited by certain inherent factors. The most apparent is sample size, which limited both the toxicity analysis of QUASAR2 and the prognosis analysis of VICTOR both by setting a minimum MAF threshold of between 1 and 5% based on statistical power and by having too few patients for the accurate estimation of SNP-level phenotypic heritability, despite there being approximately 1,000 patients in each. As genotyping/sequencing costs continue to fall, clinical trials protocols will increasingly call for the careful blood collection and genetic analysis of all patients, which will complement new collaborative efforts such those at the UK Biobank. The deluge of genetic, functional and clinical data also raises new concerns about multiple testing, which was carefully accounted for in this thesis. As it becomes customary for most genetic studies to include millions of tests, it will be necessary that reported results adjust accordingly, and that researchers in the field assume that such testing has been performed – otherwise false positives will prevail. Finally, as collaborative efforts expand and datasets are analysed and re-analysed by additional groups with additional hypothesis, the value of each datapoint will increase. Although not substantiated by the results of this thesis, much time was devoted in all chapters to the re-analysis of data upon revision from collaborating entities. As collaborations become further-reaching, accurate sharing of data will precipitate new findings.

Other factors also limited the analyses of this thesis. In the meta-analysis of Chapter 3, the lack of per-patient data (ie, the use of counts) prohibited the use of covariates, despite their minimising AIC in model selection for the candidate gene and GWAS toxicity analyses of Chapter 4 and Chapter 5. One solution might have been to request covariate-adjusted summary statistics from each author – though inappropriate covariate use might have proven equally distortionary. For the study of

prognosis in VICTOR, in addition to the study design aspects mentioned for germline genetic analysis, existing statistical models were limited in their ability to incorporate information from multiple datasets. A similar consideration was whether SNP associations with toxicity in one 5-FU-based regimen could be used to inform an analysis of the same SNP (or others in the same gene) for another 5-FU-based regimen. Finally, population differences within the somewhat broad "CEU" classification may have hindered replication of QUASAR2 results by the SCP study, and may continue to do so for rare variants that tag nearby functional variation.

Despite these challenges, the results of this thesis were a promising step toward positively impacting patient care, even though a clinically useful companion diagnostic panel is not yet ready for either capecitabine toxicity or CRC prognosis; likewise, the results of this thesis should better inform use of the currently offered kits. For toxicity, the six strongly supported variants discussed above are very likely generalisable to white patients in the UK and likely applicable to white patients in Europe and North America, due to the ethnicity of included patients. The sensitivity and specificity with which this panel can predict dose limiting levels of toxicity remains sub-optimal, but it was promising to note that the two patients from the QUASAR2 trial who experienced drug-related death held rare *DPYD* variants. If screening is deemed cost effective, future such patients could be withheld 5-FU-based drugs or given a lower dose alongside prophylactic interventions such as GCSF (granulocyte colony-stimulating factor) in order to avoid neutropaenia. However, such extreme toxicity is rare and the main outstanding clinical question is how to incorporate genetic analysis into the complex phenotype that is grade 3/4 toxicity. Because genetic results presented here did not instigate clinical action in QUASAR2, we do not know whether modest dose reduction at the start of treatment can effectively prevent high toxicity. Further, we do not know whether such lowered doses reduce drug efficacy. A trial addressing these questions would be useful, as will be the survival times of

QUASAR2 patients, which will help determine whether protective toxicity variants are limited by a comparable reduction to efficacy.

Future work could add to mine. As noted above, the following would be useful: an analysis of QUASAR2 survival in terms of both prognosis and toxicity SNPs, validation of the toxicity results in additional UK-based capecitabine-treated populations and a larger sample size estimation of the heritability of 5-FU/capecitabine toxicity and CRC prognosis. Although not accurately estimated, it is probable that additional genetic variation exists that impacts both toxicity and prognosis – further analyses could identify new variants, but an estimation of the proportion of variation explained by known variants is difficult without a heritability estimate (214). For toxicity, where a densely genotyped genome-wide analysis has been performed down to the 1% frequency level, and where exonic variation has been studied down to far lower, a larger sample size is probably needed, particularly for variants with both an odds ratio below 1.5 and MAF below 10%; however, a broader sequencing effort in candidate genes might also help identify further rare variants. For prognosis, genotyping has been less dense, but additional SNP genotyping/imputation is probably a second priority behind adding events to the analysis (except as validation of associations found in QUASAR2). Given the collection of non-genetic factors affecting toxicity and prognosis, it is plausible that the impact of germline genetic variation is only modest. A further outstanding question is in regards to the mechanisms of toxicity, both related to genetic variation and to other inter-patient differences. Given the non-obvious mechanistic relationship between *TYMS* levels/activity and 5-FU-related toxicity, it will be particularly interesting to see how *ENOSF1* variation might act to cause adverse events. To start, it would be interesting to see whether rs2612091 variation affects *TYMS* activity.

In summary, for both capecitabine-related toxicity and CRC prognosis, clinical and genetic data from multiple institutes and trials have been used to identify new

germline markers and provide shortlists of variants for further study.

Appendix A

Power Calculations for 5-FU/Capecitabine Association Studies

Table A.1: Number of patients needed to achieve 90% power for detection of a SNP association with an odds ratio of 1.5 or 0.67, at different allele frequencies, given observed incidence of each toxicity phenotype in each 5-FU-based regimen. Power calculations are only shown for phenotypes with at least 5% incidence in a given regimen.

MAF (%)	Phenotype	Incidence (%)	OR	OR	Regimen	Case	Control	n Cases	n Controls	**Sample Size**
1	GLOBAL	33.8	1.5	0.67	CAP	3.38	6.62	3850	7,537	**11,387**
5	GLOBAL	33.8	1.5	0.67	CAP	3.38	6.62	815	1,596	**2,411**
10	GLOBAL	33.8	1.5	0.67	CAP	3.38	6.62	440	861	**1,301**
20	GLOBAL	33.8	1.5	0.67	CAP	3.38	6.62	265	519	**784**
40	GLOBAL	33.8	1.5	0.67	CAP	3.38	6.62	195	382	**577**
1	GLOBAL	11.8	1.5	0.67	INF	1.18	8.82	2600	19,415	**22,015**
5	GLOBAL	11.8	1.5	0.67	INF	1.18	8.82	562	4,197	**4,759**
10	GLOBAL	11.8	1.5	0.67	INF	1.18	8.82	310	2,315	**2,625**
20	GLOBAL	11.8	1.5	0.67	INF	1.18	8.82	188	1,404	**1,592**
40	GLOBAL	11.8	1.5	0.67	INF	1.18	8.82	145	1,083	**1,228**
1	GLOBAL	28.9	1.5	0.67	BOL	2.89	7.11	3500	8,615	**12,115**
5	GLOBAL	28.9	1.5	0.67	BOL	2.89	7.11	750	1,846	**2,596**
10	GLOBAL	28.9	1.5	0.67	BOL	2.89	7.11	405	997	**1,402**
20	GLOBAL	28.9	1.5	0.67	BOL	2.89	7.11	245	603	**848**
40	GLOBAL	28.9	1.5	0.67	BOL	2.89	7.11	182	448	**630**
1	DIARRHOEA	10.0	1.5	0.67	CAP	1.00	9.00	2550	22,950	**25,500**
5	DIARRHOEA	10.0	1.5	0.67	CAP	1.00	9.00	550	4,950	**5,500**
10	DIARRHOEA	10.0	1.5	0.67	CAP	1.00	9.00	300	2,700	**3,000**
20	DIARRHOEA	10.0	1.5	0.67	CAP	1.00	9.00	185	1,665	**1,850**
40	DIARRHOEA	10.0	1.5	0.67	CAP	1.00	9.00	142	1,278	**1,420**
1	DIARRHOEA	5.4	1.5	0.67	INF	0.54	9.46	2350	41,412	**43,762**
5	DIARRHOEA	5.4	1.5	0.67	INF	0.54	9.46	510	8,987	**9,497**
10	DIARRHOEA	5.4	1.5	0.67	INF	0.54	9.46	280	4,934	**5,214**
20	DIARRHOEA	5.4	1.5	0.67	INF	0.54	9.46	172	3,031	**3,203**
40	DIARRHOEA	5.4	1.5	0.67	INF	0.54	9.46	135	2,379	**2,514**
1	DIARRHOEA	12.1	1.5	0.67	BOL	1.21	8.79	2620	19,069	**21,689**
5	DIARRHOEA	12.1	1.5	0.67	BOL	1.21	8.79	562	4,090	**4,652**
10	DIARRHOEA	12.1	1.5	0.67	BOL	1.21	8.79	310	2,256	**2,566**
20	DIARRHOEA	12.1	1.5	0.67	BOL	1.21	8.79	185	1,346	**1,531**
40	DIARRHOEA	12.1	1.5	0.67	BOL	1.21	8.79	145	1,055	**1,200**
1	HFS	22.3	1.5	0.67	CAP	2.23	7.77	3100	10,776	**13,876**
5	HFS	22.3	1.5	0.67	CAP	2.23	7.77	665	2,312	**2,977**
10	HFS	22.3	1.5	0.67	CAP	2.23	7.77	365	1,269	**1,634**
20	HFS	22.3	1.5	0.67	CAP	2.23	7.77	218	758	**976**
40	HFS	22.3	1.5	0.67	CAP	2.23	7.77	166	577	**743**
1	MUC/STOM	15.1	1.5	0.67	BOL	1.51	8.49	2750	15,498	**18,248**
5	MUC/STOM	15.1	1.5	0.67	BOL	1.51	8.49	590	3,325	**3,915**
10	MUC/STOM	15.1	1.5	0.67	BOL	1.51	8.49	324	1,826	**2,150**
20	MUC/STOM	15.1	1.5	0.67	BOL	1.51	8.49	196	1,105	**1,301**
40	MUC/STOM	15.1	1.5	0.67	BOL	1.51	8.49	150	845	**995**
1	NEUT	6.9	1.5	0.67	BOL	0.69	9.32	2415	32,840	**35,255**
5	NEUT	6.9	1.5	0.67	BOL	0.69	9.32	520	7,071	**7,591**
10	NEUT	6.9	1.5	0.67	BOL	0.69	9.32	287	3,903	**4,190**
20	NEUT	6.9	1.5	0.67	BOL	0.69	9.32	175	2,380	**2,555**
40	NEUT	6.9	1.5	0.67	BOL	0.69	9.32	137	1,863	**2,000**

Table A.2: Number of patients needed to achieve 90% power for detection of a SNP association with an odds ratio of 3.0 or 0.33, at different allele frequencies, given observed incidence of each toxicity phenotype in each 5-FU-based regimen. Power calculations are only shown for phenotypes with at least 5% incidence in a given regimen.

MAF (%)	Phenotype	Incidence (%)	OR	OR	Regimen	Case	Control	n Cases	n Controls	**Sample Size**
1	GLOBAL	33.8	3	0.33	CAP	3.38	6.62	377	738	**1,115**
5	GLOBAL	33.8	3	0.33	CAP	3.38	6.62	84	164	**248**
10	GLOBAL	33.8	3	0.33	CAP	3.38	6.62	48	94	**142**
20	GLOBAL	33.8	3	0.33	CAP	3.38	6.62	32	63	**95**
40	GLOBAL	33.8	3	0.33	CAP	3.38	6.62	29	57	**86**
1	GLOBAL	11.8	3	0.33	INF	1.18	8.82	210	1,568	**1,778**
5	GLOBAL	11.8	3	0.33	INF	1.18	8.82	50	373	**423**
10	GLOBAL	11.8	3	0.33	INF	1.18	8.82	30	224	**254**
20	GLOBAL	11.8	3	0.33	INF	1.18	8.82	22	164	**186**
40	GLOBAL	11.8	3	0.33	INF	1.18	8.82	22	164	**186**
1	GLOBAL	28.9	3	0.33	BOL	2.89	7.11	330	812	**1,142**
5	GLOBAL	28.9	3	0.33	BOL	2.89	7.11	75	185	**260**
10	GLOBAL	28.9	3	0.33	BOL	2.89	7.11	43	106	**149**
20	GLOBAL	28.9	3	0.33	BOL	2.89	7.11	29	71	**100**
40	GLOBAL	28.9	3	0.33	BOL	2.89	7.11	27	66	**93**
1	DIARRHOEA	10.0	3	0.33	CAP	1.00	9.00	201	1,809	**2,010**
5	DIARRHOEA	10.0	3	0.33	CAP	1.00	9.00	47	423	**470**
10	DIARRHOEA	10.0	3	0.33	CAP	1.00	9.00	29	261	**290**
20	DIARRHOEA	10.0	3	0.33	CAP	1.00	9.00	21	189	**210**
40	DIARRHOEA	10.0	3	0.33	CAP	1.00	9.00	21	189	**210**
1	DIARRHOEA	5.4	3	0.33	INF	0.54	9.46	178	3,137	**3,315**
5	DIARRHOEA	5.4	3	0.33	INF	0.54	9.46	42	740	**782**
10	DIARRHOEA	5.4	3	0.33	INF	0.54	9.46	26	458	**484**
20	DIARRHOEA	5.4	3	0.33	INF	0.54	9.46	19	335	**354**
40	DIARRHOEA	5.4	3	0.33	INF	0.54	9.46	21	370	**391**
1	DIARRHOEA	12.1	3	0.33	BOL	1.21	8.79	212	1,543	**1,755**
5	DIARRHOEA	12.1	3	0.33	BOL	1.21	8.79	50	364	**414**
10	DIARRHOEA	12.1	3	0.33	BOL	1.21	8.79	30	218	**248**
20	DIARRHOEA	12.1	3	0.33	BOL	1.21	8.79	22	160	**182**
40	DIARRHOEA	12.1	3	0.33	BOL	1.21	8.79	22	160	**182**
1	MUC/STOM	15.1	3	0.33	BOL	1.51	8.49	230	1,296	**1,526**
5	MUC/STOM	15.1	3	0.33	BOL	1.51	8.49	53	299	**352**
10	MUC/STOM	15.1	3	0.33	BOL	1.51	8.49	32	180	**212**
20	MUC/STOM	15.1	3	0.33	BOL	1.51	8.49	23	130	**153**
40	MUC/STOM	15.1	3	0.33	BOL	1.51	8.49	23	130	**153**
1	NEUT	6.9	3	0.33	BOL	0.69	9.32	184	2,502	**2,686**
5	NEUT	6.9	3	0.33	BOL	0.69	9.32	44	598	**642**
10	NEUT	6.9	3	0.33	BOL	0.69	9.32	27	367	**394**
20	NEUT	6.9	3	0.33	BOL	0.69	9.32	20	272	**292**
40	NEUT	6.9	3	0.33	BOL	0.69	9.32	21	286	**307**
1	HFS	22.3	3	0.33	CAP	2.23	7.77	280	973	**1,253**
5	HFS	22.3	3	0.33	CAP	2.23	7.77	63	219	**282**
10	HFS	22.3	3	0.33	CAP	2.23	7.77	37	129	**166**
20	HFS	22.3	3	0.33	CAP	2.23	7.77	26	90	**116**
40	HFS	22.3	3	0.33	CAP	2.23	7.77	25	87	**112**

Appendix B

5-FU Toxicity Meta-Analysis

QUASAR2 Genotyping

Table B.1: 5-FU Toxicity Meta-Analysis QUASAR2 Genotyping. N=total number of patient genotypes derived. No=not derived from that platform; Yes=derived from that platform; N/A=Not applicable due to use of another successful genotyping method.

SNP Name	dbSNP ID	Illumina Hap370CNV samples N=485	Illumina Hap610 samples N=352	Illumina Omni2.5 N=89	In Aug 2009 release of 1KG	KasPAR/ PCR based assays	LD pruned out	N
CES2 823 (830) C/G 5'UTR	rs11075646	No	No	No	Yes	Yes	No	927
CES2 intronic SNP rs11568314	rs11568314	Imputed Info<0.95	Directly genotyped	Imputed Info>0.95	Yes	No	No	441
CES2 intronic SNP rs11568311	rs11568311	Imputed Info<0.95	Directly genotyped	Imputed Info>0.95	Yes	No	No	441
CES2 6046G>A; R270H	rs8192924	No	No	No	No	No	No	0
CES2 6320 G/A		No	No	No	No	No	No	0
CES2 intronic SNP rs2241409	rs2241409	Imputed Info<0.95	Directly genotyped	Directly genotyped	Yes		No	441
CDA -451C>T	rs532545	Imputed Info>0.95	Imputed Info>0.95	Imputed Info>0.95	Yes	No	No	927
CDA -205C>G	rs603412	Imputed Info<0.95	Imputed Info<0.95	Directly genotyped	Yes	No	No	89
CDA 5'UTR SNP rs602950	rs602950	No	No	No	Yes	No	Yes r^2=0.99; rs532545	0
CDA 943insC	rs3215400	No	No	No	No	No	No	0
CDA*2; 79A>C; K27Q	rs2072671	Directly genotyped	Directly genotyped	Directly genotyped	N/A	No	No	927
CDA 575 C/T		No	No	No	No	No	No	0
CDA 794 G/A		No	No	No	No	No	No	0
CDA 771 C/G		No	No	No	No	No	No	0
UMPS (OPRT) 638G>C (Gly213Ala)	rs1801019	Proxy rs13146 (r^2=1)	Proxy rs13146 (r^2=1)	Proxy rs13146 (r^2=1)	No	No	No	927
UMPS (OPRT) 1336A>G (Ile446Val)	rs3772809	Imputed Info<0.95	Imputed Info<0.95	Directly genotyped	Yes	No	No	89

TYMP intronic SNP rs470119	rs470119	Directly genotyped	Directly genotyped	Imputed Info>0.95	Yes	No	No	927
TYMP A324A	rs131804	N/A	N/A	N/A	N/A	No	Yes r^2=1; rs470119	0
TYMP S471L	rs11479	Imputed Info>0.95	Imputed Info>0.95	Directly genotyped	Yes	No	No	927
TYMS 5'UTR TSER 3R G/C SNP	rs2853542	No	No	No	No	No	No	0
TYMS 5'VNTR	rs45445694	No	No	No	No	Yes	No	927
TYMS 3'UTR 1494indel6b	rs16430	No	No	No	No	Yes	No	474
MTHFR 677C>T; A222V	rs1801133	Directly genotyped	Directly genotyped	Directly genotyped	N/A	No	No	927
MTHFR 1298A>C; E429A	rs1801131	Imputed Info>0.95	Directly genotyped	Directly genotyped	N/A	No	No	927
DPYD *9A; 85T>C; C29R	rs1801265	Directly genotyped	Directly genotyped	Directly genotyped	N/A	No	No	927
DPYD 496A>G; M166V	rs2297595	Imputed Info>0.95	Directly genotyped	Directly genotyped	N/A	No	No	927
DPYD 623G>A; R208Q		No	No	No	No	No	No	0
DPYD 1109delTA		No	No	No	No	No	No	0
DPYD 1236G>A; E412E	rs56038477	Imputed Info>0.95	Imputed Info>0.95	Imputed via proxy (rs281121; r^2=1)	Yes	No	No	927
DPYD *4A; 1601G>A; S534N	rs1801158	No	No	Directly genotyped	N/A	Yes	No	927
DPYD *5; 1627A>G; I543V	rs1801159	Imputed Info>0.95	Directly genotyped	Directly genotyped	Yes	No	No	927
DPYD *13; 1679T>G; I560S	rs55886062	No	No	No	No	No	No	0
DPYD *2A; IVS14+1G>A; 1905+1G>A	rs3918290	No	No	Directly genotyped	No	Yes	No	927

Appendix C

Literature Search Strategy for 5-FU Toxicity Meta-Analysis

Table C.1: For the 5-FU toxicity meta-analysis of Chapter 3, we sought to identify 5-FU toxicity variants in English language papers published up to and including 2012.

Databases:
Search term entry: PubMed (83 searches)
Studies citing collected studies: GoogleScholar (56 searches)

Glossary:
T=search term
SS=set of searches entered into database
.ab=in abstract
.ti=in title

Strategy:
1) Identification of genes containing polymorphisms investigated for an association with FU toxicity (SS1)
2) Identification of polymorphisms investigated for an association with FU toxicity (SS2)
3) Identification of additional studies investigating each identified polymorphism (SS3, SS4)

N.B. Abstract reviewed unless title clearly identified citation as irrelevant.
N.B. Results of each search were reviewed consecutively until 50 consecutive citations were discarded due to irrelevance.

Searches:
T1. 5-fluorouracil
T2. 5-FU or 5FU
T3. capecitabine

SS1: [T1 or T2 or T3 AND "toxicity"].ti/.ab (3 searches)

Gene symbols generated by SS1:
T4. CES2
T5. CDA or CDD
T6. UMPS or OPRT
T7. TYMP or TP
T8. TYMS or TS
T9. MTHFR
T10. DPYD or DPD
T11. DPYS or DHP

SS2: [T4 AND (T1 or T2 or T3 AND "toxicity")].ti/.ab (3 searches each)
... (8 iterations)
[T11 AND (T1 or T2 or T3 AND "toxicity")].ti/.ab (24 searches total)

59 Polymorphism identifiers generated by SS2 (after study filtering):
T12. CDA or CDD and -205C>G or rs603412
T13. CDA or CDD and -451C>T or rs532545
T14. CDA or CDD and *2 or 79A>C or K27Q or rs2072671
T15. CDA or CDD and 5'UTR SNP and rs602950 or rs602950
T16. CDA or CDD and 575 C/T
T17. CDA or CDD and 771 C/G
T18. CDA or CDD and 794 G/A
T19. CDA or CDD and 943insC or rs3215400
T20. CES2 and 5841 G/A or rs11568312
T21. CES2 and 6046G>A or R270H or rs8192924
T22. CES2 and 6174 G/A or rs10852434
T23. CES2 and 6320 G/A
T24. CES2 and 823 or 830 C/G 5'UTR or rs11075646
T25. CES2 and 854G/C or rs1200937
T26. CES2 and intronic SNP rs11568311 or rs11568311

T27. CES2 and intronic SNP rs11568314 or rs11568314
T28. CES2 and intronic SNP rs2241409 or rs2241409
T29. DPYD or DPD and -1590 T>C
T30. DPYD or DPD and *13 or 1679T>G or I560S or rs55886062
T31. DPYD or DPD and *2A or IVS14+1G>A or 1905+1G>A or rs3918290
T32. DPYD or DPD and *4A or 1601G>A or S534N or rs1801158
T33. DPYD or DPD and *5 or 1627A>G or I543V or rs1801159
T34. DPYD or DPD and *6 or 2194G>A or V732I or rs1801160
T35. DPYD or DPD and *9A or 85T>C or C29R or rs1801265
T36. DPYD or DPD and 1003G>T or V335L or rs72549306
T37. DPYD or DPD and 1039-1042 del TG
T38. DPYD or DPD and 1109delTA
T39. DPYD or DPD and 1156G>T or E386T or rs78060119
T40. DPYD or DPD and 1236G>A or E412E or rs56038477
T41. DPYD or DPD and 1475C>T or S492L or rs72549304
T42. DPYD or DPD and 1897delC
T43. DPYD or DPD and 257C>T or P86L
T44. DPYD or DPD and 2657G>A or R886H or rs1801267
T45. DPYD or DPD and 2846T>A or D949V or rs67376798
T46. DPYD or DPD and 2858G>C or C953S
T47. DPYD or DPD and 2933A>G or H978R or rs201268750
T48. DPYD or DPD and 295-298del TCAT
T49. DPYD or DPD and 2983G>T or V995F or rs1801268
T50. DPYD or DPD and 496A>G or M166V or rs2297595
T51. DPYD or DPD and 601A>C or S201R or rs72549308
T52. DPYD or DPD and 61C>T or R21X or rs72549310
T53. DPYD or DPD and 623G>A or R208Q
T54. DPYD or DPD and 62G>A or R21Q or rs80081766
T55. DPYD or DPD and 632A>G or Y211C or rs72549307
T56. DPYD or DPD and 703C>T or R235W or rs1801266
T57. DPYD or DPD and 812delT
T58. DPYD or DPD and IVS10-15T>C
T59. DPYD or DPD and IVS5+18G>A
T60. MTHFR and 1298A>C or E429A or rs1801131
T61. MTHFR and 1793G>A or R549Q
T62. MTHFR and 677C>T or A222V or rs1801133
T63. TYMP or TP and A324A or rs131804
T64. TYMP or TP and intronic SNP rs470119 or rs470119
T65. TYMP or TP and S471L or rs11479
T66. TYMS 3'UTR 1494indel6b or rs16430
T67. TYMS 5'UTR TSER 2R/3R or 5'VNTR or rs34743033 or rs45445694
T68. TYMS 5'UTR TSER 3R G/C SNP or rs2853542
T69. UMPS or OPRT and 1336A>G (Ile446Val) or rs3772809
T70. UMPS or OPRT and 638G>C (Gly213Ala) or rs1801019

SS3: [T12 AND (T1 or T2 or T3 AND "toxicity")].ti/.ab (1 search each)
… (56 iterations)
[T70 AND (T1 or T2 or T3 AND "toxicity")].ti/.ab (56 searches total)

SS4: The earliest identified citation for each polymorphism was searched in GoogleScholar by title, using the “cited by” option. (56 searches)

Appendix D

5-FU Toxicity Meta-Analysis 2-by-2 Table Counts of Included Data

Data from all citations were extracted into 2-by-2 tables for each patient set/treatment regimen/polymorphism combination. In some instances, there is no recessive model output because at least one HT and one LT variant allele are not present in the data. In all instances, the "*TYMS* 5'VNTR – haplotype" polymorphism was analyzed under a binary model, with output shown in the "allelic" columns. QUASAR2 data include imputed genotypes and *DPYD* 1236G>A by proxy. LT="low toxicity" (grade 0/1/2); HT="high toxicity" (grade 3/4).

Table D.1: 5-FU Toxicity Meta-Analysis 2-by-2 Table Counts of Included Data.

Study Arm	Regimen	Polymorphism	Toxicity	Reference Allele	Test Allele	n Patients Genotyped	MAF	HWE ChiSq	HWE p-value	n het HT patients	n het LT patients	n homo test HT patients	n homo test LT patients	n HT het/test patients	n LT het/test patients	n homo ref HT patients	n homo ref LT patients	Allelic OR (95% CI)	Allelic p-value	Recessive OR (95% CI)	Recessive p-value
Caronia, 2011	CAP	CDA -205C>G	DIARRHOEA	C	G	129	0.419	1.701	0.192	15	55	1	18	1	1	1	39	1.46 (0.71-3.01)	0.304	0.33 (0.04-2.62)	0.292
Caronia, 2011	CAP	CDA -205C>G	GLOBAL	C	G	129	0.419	1.701	0.192	31	99	9	10	1	1	15	25	1.21 (0.74-2)	0.451	1.14 (0.54-3.85)	0.467
Caronia, 2011	CAP	CDA -205C>G	HFS	C	G	129	0.419	1.701	0.192	20	50	9	10	1	1	12	28	1.31 (0.77-2.22)	0.320	2.42 (0.89-6.56)	0.082
Caronia, 2011	CAP	CDA -451C>T	DIARRHOEA	C	T	130	0.358	4.624	0.032	15	56	1	10	1	1	1	47	1.97 (0.95-4.08)	0.067	0.64 (0.08-5.37)	0.684
QUASAR2, 2012	CAP	CDA -451C>T	DIARRHOEA	C	T	905	0.340	4.624	0.032	42	360	10	93	1	1	44	356	0.94 (0.68-1.29)	0.690	0.9 (0.45-1.78)	0.750
Caronia, 2011	CAP	CDA -451C>T	GLOBAL	C	T	130	0.358	4.624	0.032	36	35	5	6	1	1	14	34	1.58 (0.94-2.63)	0.082	1.15 (0.33-3.98)	0.825
QUASAR2, 2012	CAP	CDA -451C>T	GLOBAL	C	T	905	0.340	4.624	0.032	126	276	32	71	1	1	134	266	0.93 (0.76-1.15)	0.510	0.94 (0.6-1.46)	0.780
Caronia, 2011	CAP	CDA -451C>T	HFS	C	T	130	0.358	4.624	0.032	25	46	5	6	1	1	11	37	1.54 (0.9-2.64)	0.116	1.92 (0.55-6.7)	0.306
QUASAR2, 2012	CAP	CDA -451C>T	HFS	C	T	905	0.340	4.624	0.032	84	318	18	85	1	1	96	304	0.83 (0.65-1.05)	0.120	0.73 (0.43-1.25)	0.250
Caronia, 2011	CAP	CDA 5'UTR SNP rs602950	DIARRHOEA	C	T	130	0.350	3.605	0.058	15	54	1	10	1	1	1	49	2.05 (0.99-4.25)	0.052	0.64 (0.08-5.37)	0.684
Caronia, 2011	CAP	CDA 5'UTR SNP rs602950	GLOBAL	C	T	130	0.350	3.605	0.058	34	35	5	6	1	1	14	34	1.46 (0.97-2.44)	0.149	1.47 (0.42-5.15)	0.544
Caronia, 2011	CAP	CDA 5'UTR SNP rs602950	HFS	C	T	130	0.350	3.605	0.058	23	46	5	6	1	1	13	37	1.39 (0.81-2.39)	0.230	2.47 (0.7-8.72)	0.160
Ribelles, 2008	CAP	CDA 575 C/T	DIARRHOEA	C	T	123	0.402	0.519	0.471	2	61	3	15	1	1	1	41	3.14 (0.92-10.74)	0.068	6.8 (1.26-36.84)	0.026
Ribelles, 2008	CAP	CDA 575 C/T	GLOBAL	C	T	123	0.402	0.519	0.471	25	38	5	13	1	1	16	26	0.86 (0.51-1.47)	0.587	0.6 (0.2-1.81)	0.365
Ribelles, 2008	CAP	CDA 575 C/T	HFS	C	T	123	0.402	0.519	0.471	10	53	3	10	1	1	10	32	0.64 (0.32-1.28)	0.211	0.53 (0.11-2.5)	0.423
Ribelles, 2008	CAP	CDA 771 C/G	DIARRHOEA	C	G	123	0.459	0.120	0.729	2	61	3	22	1	1	1	34	2.46 (0.72-8.39)	0.151	4.32 (0.82-22.85)	0.085
Ribelles, 2008	CAP	CDA 771 C/G	GLOBAL	C	G	123	0.459	0.120	0.729	26	37	8	17	1	1	14	23	0.98 (0.58-1.65)	0.945	0.74 (0.29-1.89)	0.533
Ribelles, 2008	CAP	CDA 771 C/G	HFS	C	G	123	0.459	0.120	0.729	12	51	2	23	1	1	8	27	0.62 (0.32-1.21)	0.162	0.34 (0.07-1.56)	0.165
Ribelles, 2008	CAP	CDA 794 G/A	DIARRHOEA	G	A	123	0.057	0.448	0.503	0	14	0	0	1	1	6	103	0.61 (0.03-10.8)	0.735	- - - - - - - -	- - - -
Ribelles, 2008	CAP	CDA 794 G/A	GLOBAL	G	A	123	0.057	0.448	0.503	6	8	0	0	1	1	40	69	1.27 (0.43-3.79)	0.664	- - - - - - - -	- - - -
Ribelles, 2008	CAP	CDA 794 G/A	HFS	G	A	123	0.057	0.448	0.503	2	12	0	0	1	1	20	89	0.75 (0.16-3.5)	0.718	- - - - - - - -	- - - -
Caronia, 2011	CAP	CDA 943insC	DIARRHOEA	del	ins	130	0.415	0.771	0.380	13	55	0	20	1	1	4	38	0.85 (0.41-1.79)	0.675	0.13 (0.01-2.26)	0.161
Ribelles, 2008	CAP	CDA 943insC	DIARRHOEA	del	ins	122	0.418	2.584	0.108	1	67	1	16	1	1	4	33	0.45 (0.12-1.7)	0.237	1.25 (0.14-11.41)	0.843
Caronia, 2011	CAP	CDA 943insC	GLOBAL	del	ins	130	0.415	0.771	0.380	29	39	6	14	1	1	20	22	0.74 (0.45-1.22)	0.233	0.53 (0.19-1.49)	0.231
Ribelles, 2008	CAP	CDA 943insC	GLOBAL	del	ins	122	0.418	2.584	0.108	28	40	4	13	1	1	13	24	0.89 (0.52-1.51)	0.662	0.48 (0.15-1.57)	0.226
Caronia, 2011	CAP	CDA 943insC	HFS	del	ins	130	0.415	0.771	0.380	17	51	5	15	1	1	9	23	0.59 (0.34-1.02)	0.057	0.69 (0.23-2.03)	0.496
Ribelles, 2008	CAP	CDA 943insC	HFS	del	ins	122	0.418	2.584	0.108	16	52	2	15	1	1	3	34	1.33 (0.68-2.59)	0.402	0.6 (0.13-2.86)	0.525
QUASAR2, 2012	CAP	CDA*2 79A>C	DIARRHOEA	A	C	926	0.341	0.510	0.475	44	363	11	101	1	1	42	365	1 (0.73-1.37)	0.987	0.92 (0.48-1.79)	0.81
Ribelles, 2008	CAP	CDA*2 79A>C	DIARRHOEA	A	C	121	0.388	2.653	0.103	3	63	3	12	1	1	1	40	2.5 (0.7-7.48)	0.166	4.29 (0.71-25.95)	0.113
QUASAR2, 2012	CAP	CDA*2 79A>C	GLOBAL	A	C	926	0.341	0.510	0.475	130	277	37	75	1	1	134	273	0.99 (0.84-1.21)	0.908	1.03 (0.68-1.56)	0.898
Ribelles, 2008	CAP	CDA*2 79A>C	GLOBAL	A	C	121	0.388	2.653	0.103	28	38	4	10	1	1	13	28	1.08 (0.63-1.84)	0.776	0.64 (0.19-2.19)	0.481
QUASAR2, 2012	CAP	CDA*2 79A>C	HFS	A	C	924	0.340	0.527	0.468	85	320	23	89	1	1	98	309	0.88 (0.69-1.11)	0.276	0.89 (0.55-1.45)	0.634
Ribelles, 2008	CAP	CDA*2 79A>C	HFS	A	C	121	0.388	2.653	0.103	10	56	2	12	1	1	9	32	0.75 (0.37-1.51)	0.421	0.77 (0.16-3.74)	0.748
Caronia, 2011	CAP	CDA*2; 79A>C; K27Q	DIARRHOEA	A	C	130	0.354	4.097	0.043	15	55	1	10	1	1	4	48	2.01 (0.97-4.17)	0.059	0.64 (0.08-5.37)	0.694
Caronia, 2011	CAP	CDA*2; 79A>C; K27Q	GLOBAL	A	C	130	0.354	4.097	0.043	35	35	4	7	1	1	16	33	1.32 (0.79-2.21)	0.285	1 (0.27-3.62)	0.996
Caronia, 2011	CAP	CDA*2; 79A>C; K27Q	HFS	A	C	130	0.354	4.097	0.043	24	46	4	7	1	1	13	36	1.26 (0.73-2.16)	0.405	1.67 (0.46-6.15)	0.438
Ribelles, 2008	CAP	CES2 6046G>A	DIARRHOEA	G	A	122	0.012	53.549	0.000	0	1	0	0	1	1	1	5	46.2 (3.86-553.08)	0.002	63.55 (2.31-1745.26)	0.014
Ribelles, 2008	CAP	CES2 6046G>A	GLOBAL	G	A	122	0.012	53.549	0.000	0	1	0	0	1	1	1	45	7.36 (0.93-57.53)	0.326	504 (0.2-126.44)	0.325
Ribelles, 2008	CAP	CES2 6046G>A	HFS	G	A	122	0.012	53.549	0.000	0	1	0	1	1	1	2	22	9.38 (1.03-12.49)	0.764	1.47 (0.06-37.38)	0.814
Ribelles, 2008	CAP	CES2 6320 G/A	DIARRHOEA	G	A	122	0.008	0.008	0.927	0	2	0	0	1	1	1	6	3.69 (0.17-80.99)	0.408	- - - - - - - -	- - - -
Ribelles, 2008	CAP	CES2 6320 G/A	GLOBAL	G	A	122	0.008	0.008	0.927	1	1	0	0	1	1	1	5	1.66 (0.1-26.85)	0.721	- - - - - - - -	- - - -
Ribelles, 2008	CAP	CES2 6320 G/A	HFS	G	A	122	0.008	0.008	0.927	1	1	0	0	1	1	1	5	1.66 (0.1-26.85)	0.721	- - - - - - - -	- - - -
Caronia, 2011	CAP	CES2 823C>G	DIARRHOEA	C	G	130	0.135	0.235	0.628	1	28	1	2	1	1	1	83	0.59 (0.17-2.03)	0.400	3.47 (0.3-40.48)	0.321
QUASAR2, 2012	CAP	CES2 823C>G	DIARRHOEA	C	G	881	0.076	2.203	0.138	12	118	1	1	1	1	81	668	0.97 (0.55-1.73)	0.931	8.45 (0.52-136.25)	0.132
Ribelles, 2008	CAP	CES2 823C>G	DIARRHOEA	C	G	123	0.142	0.131	0.717	2	29	0	2	1	1	4	86	1.22 (0.26-5.81)	0.804	3.55 (0.15-81.9)	0.428
Caronia, 2011	CAP	CES2 823C>G	GLOBAL	C	G	130	0.135	0.235	0.628	12	171	2	11	1	1	11	57	1.17 (0.57-2.4)	0.661	12.79 (0.25-31.6)	0.407
QUASAR2, 2012	CAP	CES2 823C>G	GLOBAL	C	G	881	0.076	2.203	0.138	44	661	2	21	1	1	241	508	1.17 (0.81-1.69)	0.405	10.41 (0.5-217.58)	0.131
Ribelles, 2008	CAP	CES2 823C>G	GLOBAL	C	G	123	0.142	0.131	0.717	13	18	1	1	1	1	32	58	1.31 (0.63-2.7)	0.472	1.69 (0.1-27.67)	0.713
Caronia, 2011	CAP	CES2 823C>G	HFS	C	G	130	0.135	0.235	0.628	11	18	1	2	1	1	29	69	1.34 (0.64-2.81)	0.444	1.09 (0.1-12.35)	0.946

| | | |

| | | |

| | | |

| | | |

| | | |

| | | |

References

[1] Abuli, A., Lozano, J. J., Rodriguez-Soler, M., Jover, R., Bessa, X., Munoz, J., Esteban-Jurado, C., Fernandez-Rozadilla, C., Carracedo, A., Ruiz-Ponte, C., Cubiella, J., Balaguer, F., Bujanda, L., Rene, J. M., Clofent, J., Morillas, J. D., Nicolas-Perez, D., Xicola, R. M., Llor, X., Pique, J. M., Andreu, M., Castells, A., Castellvi-Bel, S., and for the Gastrointestinal Oncology Group of the Spanish Gastroenterological, A. (2013). Genetic susceptibility variants associated with colorectal cancer prognosis. *Carcinogenesis*.

[2] Afzal, S., Gusella, M., Vainer, B., Vogel, U. B., Andersen, J. T., Broedbaek, K., Petersen, M., Jimenez-Solem, E., Bertolaso, L., Barile, C., Padrini, R., Pasini, F., Jensen, S. A., and Poulsen, H. E. (2011). Combinations of polymorphisms in genes involved in the 5-fluorouracil metabolism pathway are associated with gastrointestinal toxicity in chemotherapy-treated colorectal cancer patients. *Clin Cancer Res*, 17(11):3822–9.

[3] Afzal, S., Jensen, S. A., Vainer, B., Vogel, U., Matsen, J. P., Sorensen, J. B., Andersen, P. K., and Poulsen, H. E. (2009). Mthfr polymorphisms and 5-fu-based adjuvant chemotherapy in colorectal cancer. *Ann Oncol*, 20(10):1660–6.

[4] Akaike, H. (1981). Citation classic - a new look at the statistical-model identification. *Current Contents/Engineering Technology Applied Sciences*, (51):22–22.

[5] Al Khalaf, M. M., Thalib, L., and Doi, S. A. R. (2011). Combining heterogenous studies using the random-effects model is a mistake and leads to inconclusive meta-analyses. *Journal of Clinical Epidemiology*, 64(2):119–123.

[6] Altshuler, D., Daly, M. J., and Lander, E. S. (2008). Genetic mapping in human disease. *Science*, 322(5903):881–8.

[7] Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., and Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. *Nat Protoc*, 5(9):1564–73.

[8] Andreyev, H. J., Norman, A. R., Cunningham, D., Oates, J., Dix, B. R., Iacopetta, B. J., Young, J., Walsh, T., Ward, R., Hawkins, N., Beranek, M., Jandik, P., Benamouzig, R., Jullian, E., Laurent-Puig, P., Olschwang, S., Muller, O., Hoffmann, I., Rabes, H. M., Zietz, C., Troungos, C., Valavanis, C., Yuen, S. T., Ho, J. W., Croke, C. T., O'Donoghue, D. P., Giaretti, W., Rapallo, A., Russo, A., Bazan, V., Tanaka, M., Omura, K., Azuma, T., Ohkusa, T., Fujimori, T., Ono, Y., Pauly, M., Faber, C., Glaesener, R., de Goeij, A. F., Arends, J. W., Andersen, S. N., Lovig, T., Breivik, J., Gaudernack, G., Clausen, O. P., De Angelis, P. D., Meling, G. I., Rognum, T. O., Smith, R., Goh, H. S., Font, A., Rosell, R., Sun, X. F., Zhang, H., Benhattar, J., Losi, L., Lee, J. Q., Wang, S. T., Clarke, P. A., Bell, S., Quirke, P., Bubb, V. J., Piris, J., Cruickshank, N. R., Morton, D., Fox, J. C., Al-Mulla, F., Lees, N., Hall, C. N., Snary, D., Wilkinson, K., Dillon, D., Costa, J., Pricolo, V. E., Finkelstein, S. D., Thebo, J. S., Senagore, A. J., Halter, S. A., Wadler, S., Malik, S., Krtolica, K., and Urosevic, N. (2001). Kirsten ras mutations in patients with colorectal cancer: the 'rascal ii' study. *Br J Cancer*, 85(5):692–6.

[9] Arias, S., Penchaszadeh, V. B., Pintocisternas, J., and Larrauri, S. (1980). The ivic syndrome - a new autosomal dominant complex pleiotropic syndrome with radial ray hypoplasia, hearing impairment, external ophthalmoplegia, and thrombocytopenia. *American Journal of Medical Genetics*, 6(1):25–59.

[10] Azzato, E. M., Greenberg, D., Shah, M., Blows, F., Driver, K. E., Caporaso, N. E., and Pharoah, P. D. P. (2009). Prevalent cases in observational studies of cancer survival: do they bias hazard ratio estimates? *British Journal of Cancer*, 100(11):1806–1811.

[11] Azzato, E. M., Pharoah, P. D., Harrington, P., Easton, D. F., Greenberg, D., Capo-

raso, N. E., Chanock, S. J., Hoover, R. N., Thomas, G., Hunter, D. J., and Kraft, P. (2010). A genome-wide association study of prognosis in breast cancer. *Cancer Epidemiol Biomarkers Prev*, 19(4):1140–3.

[12] Bailar, J. C., r. (1997). The promise and problems of meta-analysis. *N Engl J Med*, 337(8):559–61.

[13] Bakker, P. R., Bakker, E., Amin, N., van Duijn, C. M., van Os, J., and van Harten, P. N. (2012). Candidate gene-based association study of antipsychotic-induced movement disorders in long-stay psychiatric patients: A prospective study. *Plos One*, 7(5).

[14] Baldwin, R. M., Owzar, K., Zembutsu, H., Chhibber, A., Kubo, M., Jiang, C., Watson, D., Eclov, R. J., Mefford, J., McLeod, H. L., Friedman, P. N., Hudis, C. A., Winer, E. P., Jorgenson, E. M., Witte, J. S., Shulman, L. N., Nakamura, Y., Ratain, M. J., and Kroetz, D. L. (2012). A genome-wide association study identifies novel loci for paclitaxel-induced sensory peripheral neuropathy in calgb 40101. *Clinical Cancer Research*, 18(18):5099–5109.

[15] Balzer, S., Malde, K., Lanzen, A., Sharma, A., and Jonassen, I. (2010). Characteristics of 454 pyrosequencing data-enabling realistic simulation with flowsim. *Bioinformatics*, 26(18):i420–i425.

[16] Bayraktar, S., Thompson, P. A., Yoo, S. Y., Do, K. A., Sahin, A. A., Arun, B. K., Bondy, M. L., and Brewster, A. M. (2013). The relationship between eight gwas-identified single-nucleotide polymorphisms and primary breast cancer outcomes. *Oncologist*, 18(5):493–500.

[17] Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological*, 57(1):289–300.

[18] Berman, N. G. and Parker, R. A. (2002). Meta-analysis: neither quick nor easy. *BMC Med Res Methodol*, 2:10.

[19] Boige, V., Mendiboure, J., Pignon, J. P., Loriot, M. A., Castaing, M., Barrois, M., Malka, D., Tregouet, D. A., Bouche, O., Le Corre, D., Miran, I., Mulot, C., Ducreux, M., Beaune, P., and Laurent-Puig, P. (2010). Pharmacogenetic assessment of toxicity and outcome in patients with metastatic colorectal cancer treated with lv5fu2, folfox, and folfiri: Ffcd 2000-05. *J Clin Oncol*, 28(15):2556–64.

[20] Boisdron-Celle, M., Remaud, G., Traore, S., Poirier, A. L., Gamelin, L., Morel, A., and Gamelin, E. (2007). 5-fluorouracil-related severe toxicity: a comparison of different methods for the pretherapeutic detection of dihydropyrimidine dehydrogenase deficiency. *Cancer Lett*, 249(2):271–82.

[21] Bolton, K. L., Chenevix-Trench, G., Goh, C., Sadetzki, S., Ramus, S. J., Karlan, B. Y., Lambrechts, D., Despierre, E., Barrowdale, D., McGuffog, L., Healey, S., Easton, D. F., Sinilnikova, O., Benitez, J., Garcia, M. J., Neuhausen, S., Gail, M. H., Hartge, P., Peock, S., Frost, D., Evans, G., Eeles, R., Godwin, A. K., Daly, M. B., Kwong, A., Ma, E. S. K., Lazaro, C., Blanco, I., Montagna, M., D'Andrea, E., Nicoletto, M. O., Johnatty, S. E., Kruger, S., Jensen, A., Hogdall, E., Goode, E. L., Fridley, B. L., Loud, J. T., Greene, M. H., Mai, P. L., Chetrit, A., Lubin, F., Hirsh-Yechezkel, G., Glendon, G., Andrulis, I. L., Toland, A. E., Senter, L., Gore, M. E., Gourley, C., Michie, C. O., Song, H. L., Tyrer, J., Whittemore, A. S., McGuire, V., Sieh, W., Kristoffersson, U., Olsson, H., Borg, A., Levine, D. A., Steele, L., Beattie, M. S., Chan, S., Nussbaum, R. L., Moysich, K. B., Gross, J., Cass, I., Walsh, C., Li, A. J., Leuchter, R., Gordon, O., Garcia-Closas, M., Gayther, S. A., Chanock, S. J., Antoniou, A. C., Pharoah, P. D. P., EMBRACE, Investigators, K., and Network, C. G. A. R. (2012). Association between brca1 and brca2 mutations and survival in women with invasive epithelial ovarian cancer. *Jama-Journal of the American Medical Association*, 307(4):382–390.

[22] Bordoni, R., Bonnal, R., Rizzi, E., Carrera, P., Benedetti, S., Cremonesi, L., Stenirri, S., Colombo, A., Montrasio, C., Bonalumi, S., Albertini, A., Bernardi, L. R., Ferrari, M., and De Bellis, G. (2008). Evaluation of human gene variant detection in amplicon pools by the gs-flx parallel pyrosequencer. *BMC Genomics*, 9:464.

[23] Bosker, F. J., Hartman, C. A., Nolte, I. M., Prins, B. P., Terpstra, P., Posthuma, D., van Veen, T., Willemsen, G., DeRijk, R. H., de Geus, E. J., Hoogendijk, W. J., Sullivan, P. F., Penninx, B. W., Boomsma, D. I., Snieder, H., and Nolen, W. A. (2011). Poor replication of candidate genes for major depressive disorder using genome-wide association data. *Molecular Psychiatry*, 16(5):516–532.

[24] Branda, R. F., Chen, Z., Brooks, E. M., Naud, S. J., Trainer, T. D., and McCormack, J. J. (2002). Diet modulates the toxicity of cancer chemotherapy in rats. *J Lab Clin Med*, 140(5):358–68.

[25] Braun, M. S., Richman, S. D., Thompson, L., Daly, C. L., Meade, A. M., Adlard, J. W., Allan, J. M., Parmar, M. K., Quirke, P., and Seymour, M. T. (2009). Association of molecular markers with toxicity outcomes in a randomized trial of chemotherapy for advanced colorectal cancer: the focus trial. *J Clin Oncol*, 27(33):5519–28.

[26] Breslow, N. and Crowley, J. (1974). Large sample study of life table and product limit estimates under random censorship. *Annals of Statistics*, 2(3):437–453.

[27] Broderick, P., Carvajal-Carmona, L., Pittman, A. M., Webb, E., Howarth, K., Rowan, A., Lubbe, S., Spain, S., Sullivan, K., Fielding, S., Jaeger, E., Vijayakrishnan, J., Kemp, Z., Gorman, M., Chandler, I., Papaemmanuil, E., Penegar, S., Wood, W., Sellick, G., Qureshi, M., Teixeira, A., Domingo, E., Barclay, E., Martin, L., Sieber, O., Kerr, D., Gray, R., Peto, J., Cazier, J. B., Tomlinson, I., and Houlston, R. S. (2007). A genome-wide association study shows that common alleles of smad7 influence colorectal cancer risk. *Nature Genetics*, 39(11):1315–1317.

[28] Browning, S. R. and Browning, B. L. (2011). Population structure can inflate snp-based heritability estimates. *Am J Hum Genet*, 89(1):191–3; author reply 193–5.

[29] Cancer Genome Atlas, N. (2012). Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, 487(7407):330–7.

[30] Capitain, O., Boisdron-Celle, M., Poirier, A. L., Abadie-Lacourtoisie, S., Morel, A.,

and Gamelin, E. (2008). The influence of fluorouracil outcome parameters on tolerance and efficacy in patients with advanced colorectal cancer. *Pharmacogenomics J*, 8(4):256–67.

[31] Caronia, D., Martin, M., Sastre, J., de la Torre, J., Garcia-Saenz, J. A., Alonso, M. R., Moreno, L. T., Pita, G., Diaz-Rubio, E., Benitez, J., and Gonzalez-Neira, A. (2011). A polymorphism in the cytidine deaminase promoter predicts severe capecitabine-induced hand-foot syndrome. *Clin Cancer Res*, 17(7):2006–13.

[32] Cassidy, J., Tabernero, J., Twelves, C., Brunet, R., Butts, C., Conroy, T., Debraud, F., Figer, A., Grossmann, J., Sawada, N., Schoffski, P., Sobrero, A., Van Cutsem, E., and Diaz-Rubio, E. (2004). Xelox (capecitabine plus oxaliplatin): active first-line therapy for patients with metastatic colorectal cancer. *J Clin Oncol*, 22(11):2084–91.

[33] Cassidy, J., Twelves, C., Van Cutsem, E., Hoff, P., Bajetta, E., Boyer, M., Bugat, R., Burger, U., Garin, A., Graeven, U., McKendric, J., Maroun, J., Marshall, J., Osterwalder, B., Perez-Manga, G., Rosso, R., Rougier, P., Schilsky, R. L., and Capecitabine Colorectal Cancer Study, G. (2002). First-line oral capecitabine therapy in metastatic colorectal cancer: a favorable safety profile compared with intravenous 5-fluorouracil/leucovorin. *Ann Oncol*, 13(4):566–75.

[34] Cecchin, E., Corona, G., Masier, S., Biason, P., Cattarossi, G., Frustaci, S., Buonadonna, A., Colussi, A., and Toffoli, G. (2005). Carboxylesterase isoform 2 mrna expression in peripheral blood mononuclear cells is a predictive marker of the irinotecan to sn38 activation step in colorectal cancer patients. *Clin Cancer Res*, 11(19 Pt 1):6901–7.

[35] Chalmers, T. C. (1991). Problems induced by meta-analyses. *Stat Med*, 10(6):971–9; discussion 979–80.

[36] Chu, J. X. and Dolnick, B. J. (2002). Natural antisense (rts alpha) rna induces site-specific cleavage of thymidylate synthase mrna. *Biochimica Et Biophysica Acta-Molecular Basis of Disease*, 1587(2-3):183–193.

[37] Chua, W., Goldstein, D., Lee, C. K., Dhillon, H., Michael, M., Mitchell, P., Clarke, S. J., and Iacopetta, B. (2009). Molecular markers of response and toxicity to folfox chemotherapy in metastatic colorectal cancer. *Br J Cancer*, 101(6):998–1004.

[38] Cicek, M. S., Slager, S. L., Achenbach, S. J., French, A. J., Blair, H. E., Fink, S. R., Foster, N. R., Kabat, B. F., Halling, K. C., Cunningham, J. M., Cerhan, J. R., Jenkins, R. B., Boardman, L. A., Petersen, G. M., Sargent, D. J., Alberts, S. R., Limburg, P. J., and Thibodeau, S. N. (2009). Functional and clinical significance of variants localized to 8q24 in colon cancer. *Cancer Epidemiol Biomarkers Prev*, 18(9):2492–500.

[39] Cohen, V., Panet-Raymond, V., Sabbaghian, N., Morin, I., Batist, G., and Rozen, R. (2003). Methylenetetrahydrofolate reductase polymorphism in advanced colorectal cancer: a novel genomic predictor of clinical response to fluoropyrimidine-based chemotherapy. *Clin Cancer Res*, 9(5):1611–5.

[40] Consortium, E. P., Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., Epstein, C. B., Frietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B. R., Landt, S. G., Lee, B. K., Pauli, F., Rosenbloom, K. R., Sabo, P., Safi, A., Sanyal, A., Shoresh, N., Simon, J. M., Song, L., Trinklein, N. D., Altshuler, R. C., Birney, E., Brown, J. B., Cheng, C., Djebali, S., Dong, X., Dunham, I., Ernst, J., Furey, T. S., Gerstein, M., Giardine, B., Greven, M., Hardison, R. C., Harris, R. S., Herrero, J., Hoffman, M. M., Iyer, S., Kelllis, M., Khatun, J., Kheradpour, P., Kundaje, A., Lassmann, T., Li, Q., Lin, X., Marinov, G. K., Merkel, A., Mortazavi, A., Parker, S. C., Reddy, T. E., Rozowsky, J., Schlesinger, F., Thurman, R. E., Wang, J., Ward, L. D., Whitfield, T. W., Wilder, S. P., Wu, W., Xi, H. S., Yip, K. Y., Zhuang, J., Bernstein, B. E., Birney, E., Dunham, I., Green, E. D., Gunter, C., Snyder, M., Pazin, M. J., Lowdon, R. F., Dillon, L. A., Adams, L. B., Kelly, C. J., Zhang, J., Wexler, J. R., Green, E. D., Good, P. J., Feingold, E. A., Bernstein, B. E., Birney, E., Crawford, G. E., Dekker, J., Elinitski, L., Farnham, P. J., Gerstein, M., Giddings, M. C., Gingeras, T. R., Green, E. D., Guigo, R., Hardison, R. C., Hubbard, T. J., Kellis, M., Kent, W. J., Lieb, J. D., Margulies, E. H., Myers, R. M.,

Snyder, M., Starnatoyannopoulos, J. A., *et al.* (2012). An integrated encyclopedia of dna elements in the human genome. *Nature*, 489(7414):57–74.

[41] Cooper, G. M., Johnson, J. A., Langaee, T. Y., Feng, H., Stanaway, I. B., Schwarz, U. I., Ritchie, M. D., Stein, C. M., Roden, D. M., Smith, J. D., Veenstra, D. L., Rettie, A. E., and Rieder, M. J. (2008). A genome-wide scan for common genetic variants with a large influence on warfarin maintenance dose. *Blood*, 112(4):1022–1027.

[42] Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 34(2):187–.

[43] Cunningham, D., Pyrhonen, S., James, R. D., Punt, C. J., Hickish, T. F., Heikkila, R., Johannesen, T. B., Starkhammar, H., Topham, C. A., Awad, L., Jacques, C., and Herait, P. (1998). Randomised trial of irinotecan plus supportive care versus supportive care alone after fluorouracil failure for patients with metastatic colorectal cancer. *Lancet*, 352(9138):1413–8.

[44] Daly, A. K. and Day, C. P. (2001). Candidate gene case-control association studies: advantages and potential pitfalls. *Br J Clin Pharmacol*, 52(5):489–99.

[45] de Gramont, A., Bosset, J. F., Milan, C., Rougier, P., Bouche, O., Etienne, P. L., Morvan, F., Louvet, C., Guillot, T., Francois, E., and Bedenne, L. (1997). Randomized trial comparing monthly low-dose leucovorin and fluorouracil bolus with bimonthly high-dose leucovorin and fluorouracil bolus plus continuous infusion for advanced colorectal cancer: a french intergroup study. *J Clin Oncol*, 15(2):808–15.

[46] de Gramont, A., Figer, A., Seymour, M., Homerin, M., Hmissi, A., Cassidy, J., Boni, C., Cortes-Funes, H., Cervantes, A., Freyer, G., Papamichael, D., Le Bail, N., Louvet, C., Hendler, D., de Braud, F., Wilson, C., Morvan, F., and Bonetti, A. (2000). Leucovorin and fluorouracil with or without oxaliplatin as first-line treatment in advanced colorectal cancer. *J Clin Oncol*, 18(16):2938–47.

[47] Deenen, M. J., Tol, J., Burylo, A. M., Doodeman, V. D., de Boer, A., Vincent, A.,

Guchelaar, H. J., Smits, P. H., Beijnen, J. H., Punt, C. J., Schellens, J. H., and Cats, A. (2011). Relationship between single nucleotide polymorphisms and haplotypes in dpyd and toxicity and efficacy of capecitabine in advanced colorectal cancer. *Clin Cancer Res*, 17(10):3455–68.

[48] Delaneau, O., Zagury, J. F., and Marchini, J. (2013). Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods*, 10(1):5–6.

[49] DerSimonian, R. and Laird, N. (1986). Meta-analysis in clinical trials. *Control Clin Trials*, 7(3):177–88.

[50] Derwinger, K., Wettergren, Y., Odin, E., Carlsson, G., and Gustavsson, B. (2009). A study of the mthfr gene polymorphism c677t in colorectal cancer. *Clin Colorectal Cancer*, 8(1):43–8.

[51] Desmoulin, F., Gilard, V., Malet-Martino, M., and Martino, R. (2002). Metabolism of capecitabine, an oral fluorouracil prodrug: F-19 nmr studies in animal models and human urine. *Drug Metabolism and Disposition*, 30(11):1221–1229.

[52] Dimas, A. S., Deutsch, S., Stranger, B. E., Montgomery, S. B., Borel, C., Attar-Cohen, H., Ingle, C., Beazley, C., Arcelus, M. G., Sekowska, M., Gagnebin, M., Nisbett, J., Deloukas, P., Dermitzakis, E. T., and Antonarakis, S. E. (2009). Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science*, 325(5945):1246–1250.

[53] Dolnick, B. J. (1993). Cloning and characterization of a naturally-occurring antisense rna to human thymidylate synthase messenger-rna. *Nucleic Acids Research*, 21(8):1747–1752.

[54] Dolnick, B. J., Angelino, N. J., Dolnick, R., and Sufrin, J. R. (2003). A novel function for the rts gene. *Cancer Biology Therapy*, 2(4):364–369.

[55] Dolnick, B. J. and Black, A. R. (1996). Alternate splicing of the rts gene product and its overexpression in a 5-fluorouracil-resistant cell line. *Cancer Res*, 56(14):3207–10.

[56] Domingo, E., Ramamoorthy, R., Oukrif, D., Rosmarin, D., Presz, M., Wang, H., Pulker, H., Lockstone, H., Hveem, T., Cranston, T., Danielsen, H., Novelli, M., Davidson, B., Xu, Z. Z., Molloy, P., Johnstone, E., Holmes, C., Midgley, R., Kerr, D., Sieber, O., and Tomlinson, I. (2013). Use of multivariate analysis to suggest a new molecular classification of colorectal cancer. *J Pathol*, 229(3):441–8.

[57] Dotor, E., Cuatrecases, M., Martinez-Iniesta, M., Navarro, M., Vilardell, F., Guino, E., Pareja, L., Figueras, A., Mollevi, D. G., Serrano, T., de Oca, J., Peinado, M. A., Moreno, V., Germa, J. R., Capella, G., and Villanueva, A. (2006). Tumor thymidylate synthase 1494de16 genotype as a prognostic factor in colorectal cancer patients receiving fluorouracil-based adjuvant treatment. *Journal of Clinical Oncology*, 24(10):1603–1611.

[58] Douillard, J. Y., Cunningham, D., Roth, A. D., Navarro, M., James, R. D., Karasek, P., Jandik, P., Iveson, T., Carmichael, J., Alakl, M., Gruia, G., Awad, L., and Rougier, P. (2000). Irinotecan combined with fluorouracil compared with fluorouracil alone as first-line treatment for metastatic colorectal cancer: a multicentre randomised trial. *Lancet*, 355(9209):1041–7.

[59] Dunlop, M. G., Dobbins, S. E., Farrington, S. M., Jones, A. M., Palles, C., Whiffin, N., Tenesa, A., Spain, S., Broderick, P., Ooi, L. Y., Domingo, E., Smillie, C., Henrion, M., Frampton, M., Martin, L., Grimes, G., Gorman, M., Semple, C., Ma, Y. P., Barclay, E., Prendergast, J., Cazier, J. B., Olver, B., Penegar, S., Lubbe, S., Chander, I., Carvajal-Carmona, L. G., Ballereau, S., Lloyd, A., Vijayakrishnan, J., Zgaga, L., Rudan, I., Theodoratou, E., Colorectal Tumour Gene Identification, C., Starr, J. M., Deary, I., Kirac, I., Kovacevic, D., Aaltonen, L. A., Renkonen-Sinisalo, L., Mecklin, J. P., Matsuda, K., Nakamura, Y., Okada, Y., Gallinger, S., Duggan, D. J., Conti, D., Newcomb, P., Hopper, J., Jenkins, M. A., Schumacher, F., Casey, G., Easton, D., Shah, M., Pharoah, P., Lindblom, A., Liu, T., Swedish Low-Risk Colorectal Cancer Study, G., Smith, C. G., West, H., Cheadle, J. P., Group, C. C., Midgley, R., Kerr, D. J., Campbell, H., Tomlinson, I. P., and Houlston, R. S. (2012a). Common variation near cdkn1a, pold3 and shroom2 influences colorectal cancer risk. *Nat Genet*, 44(7):770–6.

[60] Dunlop, M. G., Dobbins, S. E., Farrington, S. M., Jones, A. M., Palles, C., Whiffin, N., Tenesa, A., Spain, S., Broderick, P., Ooi, L. Y., Domingo, E., Smillie, C., Henrion, M., Frampton, M., Martin, L., Grimes, G., Gorman, M., Semple, C., Ma, Y. P., Barclay, E., Prendergast, J., Cazier, J. B., Olver, B., Penegar, S., Lubbe, S., Chander, I., Carvajal-Carmona, L. G., Ballereau, S., Lloyd, A., Vijayakrishnan, J., Zgaga, L., Rudan, I., Theodoratou, E., Colorectal Tumour Gene Identification, C., Starr, J. M., Deary, I., Kirac, I., Kovacevic, D., Aaltonen, L. A., Renkonen-Sinisalo, L., Mecklin, J. P., Matsuda, K., Nakamura, Y., Okada, Y., Gallinger, S., Duggan, D. J., Conti, D., Newcomb, P., Hopper, J., Jenkins, M. A., Schumacher, F., Casey, G., Easton, D., Shah, M., Pharoah, P., Lindblom, A., Liu, T., Swedish Low-Risk Colorectal Cancer Study, G., Smith, C. G., West, H., Cheadle, J. P., Group, C. C., Midgley, R., Kerr, D. J., Campbell, H., Tomlinson, I. P., and Houlston, R. S. (2012b). Common variation near cdkn1a, pold3 and shroom2 influences colorectal cancer risk. *Nat Genet*, 44(7):770–6.

[61] Egger, M., Smith, G. D., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109):629–634.

[62] Etienne-Grimaldi, M. C., Milano, G., Maindrault-Goebel, F., Chibaudel, B., Formento, J. L., Francoual, M., Lledo, G., Andre, T., Mabro, M., Mineur, L., Flesch, M., Carola, E., and de Gramont, A. (2010). Methylenetetrahydrofolate reductase (mthfr) gene polymorphisms and folfox response in colorectal cancer patients. *Br J Clin Pharmacol*, 69(1):58–66.

[63] Fernandez-Rozadilla, C., Cazier, J. B., Moreno, V., Crous-Bou, M., Guino, E., Duran, G., Lamas, M. J., Lopez, R., Candamio, S., Gallardo, E., Pare, L., Baiget, M., Paez, D., Lopez-Fernandez, L. A., Cortejoso, L., Garcia, M. I., Bujanda, L., Gonzalez, D., Gonzalo, V., Rodrigo, L., Rene, J. M., Jover, R., Brea-Fernandez, A., Andreu, M., Bessa, X., Llor, X., Xicola, R., Palles, C., Tomlinson, I., Castellvi-Bel, S., Castells, A., Ruiz-Ponte, C., and Carracedo, A. (2012). Pharmacogenomics in colorectal cancer: a genome-wide association study to predict toxicity after 5-fluorouracil or folfox administration. *Pharmacogenomics J.*

[64] Fijneman, R. J. A., Anderson, R. A., Richards, E., Liu, J. M., Tijssen, M., Meijer, G. A., Anderson, J., Rod, A., O'Sullivan, M. G., Scott, P. M., and Cormier, R. T. (2012). Runx1 is a tumor suppressor gene in the mouse gastrointestinal tract. *Cancer Science*, 103(3):593–599.

[65] Formica, V., Palmirotta, R., Del Monte, G., Savonarola, A., Ludovici, G., De Marchis, M. L., Grenga, I., Schirru, M., Guadagni, F., and Roselli, M. (2011). Predictive value of vegf gene polymorphisms for metastatic colorectal cancer patients receiving first-line treatment including fluorouracil, irinotecan, and bevacizumab. *Int J Colorectal Dis*, 26(2):143–51.

[66] Friboulet, L., Olaussen, K. A., Pignon, J. P., Shepherd, F. A., Tsao, M. S., Graziano, S., Kratzke, R., Douillard, J. Y., Seymour, L., Pirker, R., Filipits, M., Andre, F., Solary, E., Ponsonnailles, F., Robin, A., Stoclin, A., Dorvault, N., Commo, F., Adam, J., Vanhecke, E., Saulnier, P., Thomale, J., Le Chevalier, T., Dunant, A., Rousseau, V., Le Teuff, G., Brambilla, E., and Soria, J. C. (2013). Ercc1 isoform expression and dna repair in non-small-cell lung cancer. *N Engl J Med*, 368(12):1101–10.

[67] Frosst, P., Blom, H. J., Milos, R., Goyette, P., Sheppard, C. A., Matthews, R. G., Boers, G. J., den Heijer, M., Kluijtmans, L. A., van den Heuvel, L. P., and et al. (1995). A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat Genet*, 10(1):111–3.

[68] Genomes Project, C., Abecasis, G. R., Altshuler, D., Auton, A., Brooks, L. D., Durbin, R. M., Gibbs, R. A., Hurles, M. E., and McVean, G. A. (2010). A map of human genome variation from population-scale sequencing. *Nature*, 467(7319):1061–73.

[69] Ghosh, S., Winter, J. M., Patel, K., and Kern, S. E. (2011). Reexamining a proposal: thymidylate synthase 5'-untranslated region as a regulator of translation efficiency. *Cancer Biol Ther*, 12(8):750–5.

[70] Gibson, G. (2011). Rare and common variants: twenty arguments. *Nat Rev Genet*, 13(2):135–45.

[71] Glimelius, B., Garmo, H., Berglund, A., Fredriksson, L. A., Berglund, M., Kohnke, H., Bystrom, P., Sorbye, H., and Wadelius, M. (2011). Prediction of irinotecan and 5-fluorouracil toxicity and response in patients with advanced colorectal cancer. *Pharmacogenomics J*, 11(1):61–71.

[72] Goekkurt, E., Al-Batran, S. E., Hartmann, J. T., Mogck, U., Schuch, G., Kramer, M., Jaeger, E., Bokemeyer, C., Ehninger, G., and Stoehlmacher, J. (2009). Pharmacogenetic analyses of a phase iii trial in metastatic gastroesophageal adenocarcinoma with fluorouracil and leucovorin plus either oxaliplatin or cisplatin: a study of the arbeitsgemeinschaft internistische onkologie. *J Clin Oncol*, 27(17):2863–73.

[73] Goldstein, J. I., Crenshaw, A., Carey, J., Grant, G. B., Maguire, J., Fromer, M., O'Dushlaine, C., Moran, J. L., Chambert, K., Stevens, C., Swedish Schizophrenia, C., Consortium, A. A. S., Sklar, P., Hultman, C. M., Purcell, S., McCarroll, S. A., Sullivan, P. F., Daly, M. J., and Neale, B. M. (2012). zcall: a rare variant caller for array-based genotyping: genetics and population analysis. *Bioinformatics*, 28(19):2543–5.

[74] Goode, E. L., Dunning, A. M., Kuschel, B., Healey, C. S., Day, N. E., Ponder, B. A., Easton, D. F., and Pharoah, P. P. (2002). Effect of germ-line genetic variation on breast cancer survival in a population-based study. *Cancer Res*, 62(11):3052–7.

[75] Gray, R., Barnwell, J., McConkey, C., Hills, R. K., Williams, N. S., and Kerr, D. J. (2007). Adjuvant chemotherapy versus observation in patients with colorectal cancer: a randomised study. *Lancet*, 370(9604):2020–9.

[76] Grem, J. L. (2000). 5-fluorouracil: forty-plus and still ticking. a review of its preclinical and clinical development. *Invest New Drugs*, 18(4):299–313.

[77] Gross, E., Busse, B., Riemenschneider, M., Neubauer, S., Seck, K., Klein, H. G., Kiechle, M., Lordick, F., and Meindl, A. (2008). Strong association of a common dihydropyrimidine dehydrogenase gene polymorphism with fluoropyrimidine-related toxicity in cancer patients. *PLoS One*, 3(12):e4003.

[78] Gunderson, K. L., Steemers, F. J., Lee, G., Mendoza, L. G., and Chee, M. S. (2005). A genome-wide scalable snp genotyping assay using microarray technology. *Nature Genetics*, 37(5):549–554.

[79] Gusella, M., Frigo, A. C., Bolzonella, C., Marinelli, R., Barile, C., Bononi, A., Crepaldi, G., Menon, D., Stievano, L., Toso, S., Pasini, F., Ferrazzi, E., and Padrini, R. (2009). Predictors of survival and toxicity in patients on adjuvant therapy with 5-fluorouracil for colorectal cancer. *Br J Cancer*, 100(10):1549–57.

[80] Haller, D. G., Cassidy, J., Clarke, S. J., Cunningham, D., Van Cutsem, E., Hoff, P. M., Rothenberg, M. L., Saltz, L. B., Schmoll, H. J., Allegra, C., Bertino, J. R., Douillard, J. Y., Gustavsson, B. G., Milano, G., O'Connell, M., Rustum, Y., Tabernero, J., Gilberg, F., Sirzen, F., and Twelves, C. (2008). Potential regional differences for the tolerability profiles of fluoropyrimidines. *J Clin Oncol*, 26(13):2118–23.

[81] Han, J. Y., Shin, E. S., Lee, Y. S., Ghang, H. Y., Kim, S. Y., Hwang, J. A., Kim, J. Y., and Lee, J. S. (2012). A genome-wide association study for irinotecan-related severe toxicities in patients with advanced non-small-cell lung cancer. *Pharmacogenomics J.*

[82] Hardy, L. W., Finer-Moore, J. S., Montfort, W. R., Jones, M. O., Santi, D. V., and Stroud, R. M. (1987). Atomic structure of thymidylate synthase: target for rational drug design. *Science*, 235(4787):448–55.

[83] Harris, R. J., Bradburn, M. J., Deeks, J. J., Harbord, R. M., Altman, D. G., and Sterne, J. A. C. (2008). metan: fixed- and random-effects meta-analysis. *Stata Journal*, 8(1):3–28.

[84] Heggie, G. D., Sommadossi, J. P., Cross, D. S., Huster, W. J., and Diasio, R. B. (1987). Clinical pharmacokinetics of 5-fluorouracil and its metabolites in plasma, urine, and bile. *Cancer Res*, 47(8):2203–6.

[85] Higgins, J. P. and Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Stat Med*, 21(11):1539–58.

[86] Higgins, J. P., Thompson, S. G., Deeks, J. J., and Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414):557–60.

[87] Hon, G. C., Hawkins, R. D., and Ren, B. (2009). Predictive chromatin signatures in the mammalian genome. *Human Molecular Genetics*, 18:R195–R201.

[88] Horie, N., Aiba, H., Oguro, K., Hojo, H., and Takeishi, K. (1995). Functional analysis and dna polymorphism of the tandemly repeated sequences in the 5'-terminal regulatory region of the human gene for thymidylate synthase. *Cell Struct Funct*, 20(3):191–7.

[89] Houlston, R. S., Cheadle, J., Dobbins, S. E., Tenesa, A., Jones, A. M., Howarth, K., Spain, S. L., Broderick, P., Domingo, E., Farrington, S., Prendergast, J. G. D., Pittman, A. M., Theodoratou, E., Smith, C. G., Olver, B., Walther, A., Barnetson, R. A., Churchman, M., Jaeger, E. E. M., Penegar, S., Barclay, E., Martin, L., Gorman, M., Mager, R., Johnstone, E., Midgley, R., Niittymaki, I., Tuupanen, S., Colley, J., Idziaszczyk, S., Thomas, H. J. W., Lucassen, A. M., Evans, D. G. R., Maher, E. R., Maughan, T., Dimas, A., Dermitzakis, E., Cazier, J. B., Aaltonen, L. A., Pharoah, P., Kerr, D. J., Carvajal-Carmona, L. G., Campbell, H., Dunlop, M. G., Tomlinson, I. P. M., Consortium, C., Consortium, C., Grp, C. C., and Grp, C. C. (2010). Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. *Nature Genetics*, 42(11):973–U89.

[90] Houlston, R. S., Webb, E., Broderick, P., Pittman, A. M., Di Bernardo, M. C., Lubbe, S., Chandler, I., Vijayakrishnan, J., Sullivan, K., Penegar, S., Colorectal Cancer Association Study, C., Carvajal-Carmona, L., Howarth, K., Jaeger, E., Spain, S. L., Walther, A., Barclay, E., Martin, L., Gorman, M., Domingo, E., Teixeira, A. S., Co, R. G. I. C., Kerr, D., Cazier, J. B., Niittymaki, I., Tuupanen, S., Karhu, A., Aaltonen, L. A., Tomlinson, I. P., Farrington, S. M., Tenesa, A., Prendergast, J. G., Barnetson, R. A., Cetnarskyj, R., Porteous, M. E., Pharoah, P. D., Koessler, T., Hampe, J., Buch, S., Schafmayer, C., Tepel, J., Schreiber, S., Volzke, H., Chang-Claude, J., Hoffmeister, M., Brenner, H., Zanke, B. W., Montpetit, A., Hudson, T. J., Gallinger, S., International Colorectal Can-

cer Genetic Association, C., Campbell, H., and Dunlop, M. G. (2008). Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat Genet*, 40(12):1426–35.

[91] Howie, B., Fuchsberger, C., Stephens, M., Marchini, J., and Abecasis, G. R. (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet*, 44(8):955–9.

[92] Howie, B., Marchini, J., and Stephens, M. (2011). Genotype imputation with thousands of genomes. *G3 (Bethesda)*, 1(6):457–70.

[93] Howie, B. N., Donnelly, P., and Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*, 5(6):e1000529.

[94] Hu, L., Wu, C., Zhao, X., Heist, R., Su, L., Zhao, Y., Han, B., Cao, S., Chu, M., Dai, J., Dong, J., Shu, Y., Xu, L., Chen, Y., Wang, Y., Lu, F., Jiang, Y., Yu, D., Chen, H., Tan, W., Ma, H., Chen, J., Jin, G., Wu, T., Lu, D., Christiani, D. C., Lin, D., Hu, Z., and Shen, H. (2012). Genome-wide association study of prognosis in advanced non-small cell lung cancer patients receiving platinum-based chemotherapy. *Clin Cancer Res*, 18(19):5507–14.

[95] Iglesias, D., Nejda, N., Azcoita, M. M., Schwartz, S., J., Gonzalez-Aguilera, J. J., and Fernandez-Peralta, A. M. (2009). Effect of cox2 -765g>c and c.3618a>g polymorphisms on the risk and survival of sporadic colorectal cancer. *Cancer Causes Control*, 20(8):1421–9.

[96] Ingle, J. N., Schaid, D. J., Goss, P. E., Liu, M., Mushiroda, T., Chapman, J. A. W., Kubo, M., Jenkins, G. D., Batzler, A., Shepherd, L., Pater, J., Wang, L. W., Ellis, M. J., Stearns, V., Rohrer, D. C., Goetz, M. P., Pritchard, K. I., Flockhart, D. A., Nakamura, Y., and Weinshilboum, R. M. (2010). Genome-wide associations and functional genomic studies of musculoskeletal adverse events in women receiving aromatase inhibitors. *Journal of Clinical Oncology*, 28(31):4674–4682.

[97] Innocenti, F., Owzar, K., Cox, N. L., Evans, P., Kubo, M., Zembutsu, H., Jiang, C., Hollis, D., Mushiroda, T., Li, L., Friedman, P., Wang, L. W., Glubb, D., Hurwitz, H., Giacomini, K. M., McLeod, H. L., Goldberg, R. M., Schilsky, R. L., Kindler, H. L., Nakamura, Y., and Ratain, M. J. (2012). A genome-wide association study of overall survival in pancreatic cancer patients treated with gemcitabine in calgb 80303. *Clinical Cancer Research*, 18(2):577–584.

[Institute] Institute, N. C. Common terminology criteria for adverse events v4.0.

[99] Institute., N. C. Pdq® colon cancer treatment.

[100] Institute., N. C. Pdq® rectal cancer treatment.

[101] Ioannidis, J. P. A., Tarone, R., and McLaughlin, J. K. (2011). The false-positive to false-negative ratio in epidemiologic studies. *Epidemiology*, 22(4):450–456.

[102] Jaeger, E., Webb, E., Howarth, K., Carvajal-Carmona, L., Rowan, A., Broderick, P., Walther, A., Spain, S., Pittman, A., Kemp, Z., Sullivan, K., Heinimann, K., Lubbe, S., Domingo, E., Barclay, E., Martin, L., Gorman, M., Chandler, I., Vijayakrishnan, J., Wood, W., Papaemmanuil, E., Penegar, S., Qureshi, M., Consortium, C., Farrington, S., Tenesa, A., Cazier, J. B., Kerr, D., Gray, R., Peto, J., Dunlop, M., Campbell, H., Thomas, H., Houlston, R., and Tomlinson, I. (2008). Common genetic variants at the crac1 (hmps) locus on chromosome 15q13.3 influence colorectal cancer risk. *Nat Genet*, 40(1):26–8.

[103] Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics. *CA Cancer J Clin*, 61(2):69–90.

[104] Jennings, B. A., Kwok, C. S., Willis, G., Matthews, V., Wawruch, P., and Loke, Y. K. (2012). Functional polymorphisms of folate metabolism and response to chemotherapy for colorectal cancer, a systematic review and meta-analysis. *Pharmacogenet Genomics*, 22(4):290–304.

[105] Jennings, B. A., Willis, G. A., Skinner, J., and Relton, C. L. (2010). Genetic selection? a study of individual variation in the enzymes of folate metabolism. *BMC Med Genet*, 11:18.

[106] Johnson, G. C. L., Esposito, L., Barratt, B. J., Smith, A. N., Heward, J., Di Genova, G., Ueda, H., Cordell, H. J., Eaves, I. A., Dudbridge, F., Twells, R. C. J., Payne, F., Hughes, W., Nutland, S., Stevens, H., Carr, P., Tuomilehto-Wolf, E., Tuomilehto, J., Gough, S. C. L., Clayton, D. G., and Todd, J. A. (2001). Haplotype tagging for the identification of common disease genes. *Nature Genetics*, 29(2):233–237.

[107] Johnson, M. R., Yan, J., Shao, L., Albin, N., and Diasio, R. B. (1997). Semi-automated radioassay for determination of dihydropyrimidine dehydrogenase (dpd) activity. screening cancer patients for dpd deficiency, a condition associated with 5-fluorouracil toxicity. *J Chromatogr B Biomed Sci Appl*, 696(2):183–91.

[108] Johnson, R. C., Nelson, G. W., Troyer, J. L., Lautenberger, J. A., Kessing, B. D., Winkler, C. A., and O'Brien, S. J. (2010). Accounting for multiple comparisons in a genome-wide association study (gwas). *BMC Genomics*, 11:724.

[109] Jones, M. R., Chua, A. K., Mengesha, E. A., Taylor, K. D., Chen, Y. D. I., Li, X. H., Krauss, R. M., Rotter, J. I., Legro, R. S., Azziz, R., Goodarzi, M. O., and Network, R. M. (2012). Metabolic and cardiovascular genes in polycystic ovary syndrome: A candidate-wide association study (cwas). *Steroids*, 77(4):317–322.

[110] Jonker, D. J., Maroun, J. A., and Kocha, W. (2000). Survival benefit of chemotherapy in metastatic colorectal cancer: a meta-analysis of randomized controlled trials. *Br J Cancer*, 82(11):1789–94.

[111] Kaplan, E. L. and Meier, P. (1958). Nonparametric-estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481.

[112] Kawakami, K., Salonga, D., Park, J. M., Danenberg, K. D., Uetake, H., Brabender, J., Omura, K., Watanabe, G., and Danenberg, P. V. (2001). Different lengths of a poly-

morphic repeat sequence in the thymidylate synthase gene affect translational efficiency but not its gene expression. *Clin Cancer Res*, 7(12):4096–101.

[113] Kiyotani, K., Uno, S., Mushiroda, T., Takahashi, A., Kubo, M., Mitsuhata, N., Ina, S., Kihara, C., Kimura, Y., Yamaue, H., Hirata, K., Nakamura, Y., and Zembutsu, H. (2012). A genome-wide association study identifies four genetic markers for hematological toxicities in cancer patients receiving gemcitabine therapy. *Pharmacogenet Genomics*, 22(4):229–35.

[114] Koessler, T., Azzato, E. M., Perkins, B., Macinnis, R. J., Greenberg, D., Easton, D. F., and Pharoah, P. D. (2009). Common germline variation in mismatch repair genes and survival after a diagnosis of colorectal cancer. *Int J Cancer*, 124(8):1887–91.

[115] Largillier, R., Etienne-Grimaldi, M. C., Formento, J. L., Ciccolini, J., Nebbia, J. F., Ginot, A., Francoual, M., Renee, N., Ferrero, J. M., Foa, C., Namer, M., Lacarelle, B., and Milano, G. (2006). Pharmacogenetics of capecitabine in advanced breast cancer patients. *Clin Cancer Res*, 12(18):5496–502.

[116] Lecomte, T., Ferraz, J. M., Zinzindohoue, F., Loriot, M. A., Tregouet, D. A., Landi, B., Berger, A., Cugnenc, P. H., Jian, R., Beaune, P., and Laurent-Puig, P. (2004). Thymidylate synthase gene polymorphism predicts toxicity in colorectal cancer patients receiving 5-fluorouracil-based chemotherapy. *Clin Cancer Res*, 10(17):5880–8.

[117] Leichman, C. G., Lenz, H. J., Leichman, L., Danenberg, K., Baranda, J., Groshen, S., Boswell, W., Metzger, R., Tan, M., and Danenberg, P. V. (1997). Quantitation of intratumoral thymidylate synthase expression predicts for disseminated colorectal cancer response and resistance to protracted-infusion fluorouracil and weekly leucovorin. *Journal of Clinical Oncology*, 15(10):3223–3229.

[118] LeLorier, J., Gregoire, G., Benhaddad, A., Lapierre, J., and Derderian, F. (1997a). Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med*, 337(8):536–42.

[119] LeLorier, J., Gregoire, G., Benhaddad, A., Lapierre, J., and Derderian, F. (1997b). Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *New England Journal of Medicine*, 337(8):536–542.

[120] Lenz, H. J., Leichman, C. G., Danenberg, K. D., Danenberg, P. V., Groshen, S., Cohen, H., Laine, L., Crookes, P., Silberman, H., Baranda, J., Garcia, Y., Li, J., and Leichman, L. (1996). Thymidylate synthase mrna level in adenocarcinoma of the stomach: A predictor for primary tumor response and overall survival. *Journal of Clinical Oncology*, 14(1):176–182.

[121] Li, Q. Y., Seo, J. H., Stranger, B., McKenna, A., Pe'er, I., LaFramboise, T., Brown, M., Tyekucheva, S., and Freedman, M. L. (2013). Integrative eqtl-based analyses reveal the biology of breast cancer risk loci. *Cell*, 152(3):633–641.

[122] Liao, X., Lochhead, P., Nishihara, R., Morikawa, T., Kuchiba, A., Yamauchi, M., Imamura, Y., Qian, Z. R., Baba, Y., Shima, K., Sun, R., Nosho, K., Meyerhardt, J. A., Giovannucci, E., Fuchs, C. S., Chan, A. T., and Ogino, S. (2012). Aspirin use, tumor pik3ca mutation, and colorectal-cancer survival. *N Engl J Med*, 367(17):1596–606.

[123] Lichtenstein, P., Holm, N. V., Verkasalo, P. K., Iliadou, A., Kaprio, J., Koskenvuo, M., Pukkala, E., Skytthe, A., and Hemminki, K. (2000). Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from sweden, denmark, and finland. *N Engl J Med*, 343(2):78–85.

[124] Lin, T. M., Tang, N., and Ostap, E. M. (2005). Biochemical and motile properties of myo1b splice isoforms. *Journal of Biological Chemistry*, 280(50):41562–41567.

[125] Liu, M., Cao, D., Russell, R., Handschumacher, R. E., and Pizzorno, G. (1998a). Expression, characterization, and detection of human uridine phosphorylase and identification of variant uridine phosphorolytic activity in selected human tumors. *Cancer Res*, 58(23):5418–24.

[126] Liu, M., Cao, D., Russell, R., Handschumacher, R. E., and Pizzorno, G. (1998b).

Expression, characterization, and detection of human uridine phosphorylase and identification of variant uridine phosphorolytic activity in selected human tumors. *Cancer Res*, 58(23):5418–24.

[127] Longley, D. B., Harkin, D. P., and Johnston, P. G. (2003). 5-fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer*, 3(5):330–8.

[128] Louvet, C., Andre, T., Tigaud, J. M., Gamelin, E., Douillard, J. Y., Brunet, R., Francois, E., Jacob, J. H., Levoir, D., Taamma, A., Rougier, P., Cvitkovic, E., and de Gramont, A. (2002). Phase ii study of oxaliplatin, fluorouracil, and folinic acid in locally advanced or metastatic gastric cancer patients. *J Clin Oncol*, 20(23):4543–8.

[129] Lu, Z., Zhang, R., and Diasio, R. B. (1993). Dihydropyrimidine dehydrogenase activity in human peripheral blood mononuclear cells and liver: population characteristics, newly identified deficient patients, and clinical implication in 5-fluorouracil chemotherapy. *Cancer Res*, 53(22):5433–8.

[130] Macdonald, J. S. (2002). Vive la difference: sex and fluorouracil toxicity. *J Clin Oncol*, 20(6):1439–41.

[131] Mandola, M. V., Stoehlmacher, J., Muller-Weeks, S., Cesarone, G., Yu, M. C., Lenz, H. J., and Ladner, R. D. (2003). A novel single nucleotide polymorphism within the 5' tandem repeat polymorphism of the thymidylate synthase gene abolishes usf-1 binding and alters transcriptional activity. *Cancer Res*, 63(11):2898–904.

[132] Mandola, M. V., Stoehlmacher, J., Zhang, W., Groshen, S., Yu, M. C., Iqbal, S., Lenz, H. J., and Ladner, R. D. (2004). A 6 bp polymorphism in the thymidylate synthase gene causes message instability and is associated with decreased intratumoral ts mrna levels. *Pharmacogenetics*, 14(5):319–27.

[133] Marchini, J. and Howie, B. (2010). Genotype imputation for genome-wide association studies. *Nat Rev Genet*, 11(7):499–511.

[134] Marchini, J., Howie, B., Myers, S., McVean, G., and Donnelly, P. (2007a). A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*, 39(7):906–13.

[135] Marchini, J., Howie, B., Myers, S., McVean, G., and Donnelly, P. (2007b). A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*, 39(7):906–13.

[136] Marcuello, E., Altes, A., del Rio, E., Cesar, A., Menoyo, A., and Baiget, M. (2004). Single nucleotide polymorphism in the 5' tandem repeat sequences of thymidylate synthase gene predicts for response to fluorouracil-based chemotherapy in advanced colorectal cancer patients. *Int J Cancer*, 112(5):733–7.

[137] Mardis, E. R. (2008). Next-generation dna sequencing methods. *Annual Review of Genomics and Human Genetics*, 9:387–402.

[138] Martinez-Balibrea, E., Abad, A., Aranda, E., Sastre, J., Manzano, J. L., Diaz-Rubio, E., Gomez-Espana, A., Aparicio, J., Garcia, T., Maestu, I., Martinez-Cardus, A., Gines, A., and Guino, E. (2008). Pharmacogenetic approach for capecitabine or 5-fluorouracil selection to be combined with oxaliplatin as first-line chemotherapy in advanced colorectal cancer. *Eur J Cancer*, 44(9):1229–37.

[139] Martinez-Balibrea, E., Abad, A., Martinez-Cardus, A., Gines, A., Valladares, M., Navarro, M., Aranda, E., Marcuello, E., Benavides, M., Massuti, B., Carrato, A., Layos, L., Manzano, J. L., and Moreno, V. (2010). Ugt1a and tyms genetic variants predict toxicity and response of colorectal cancer patients treated with first-line irinotecan and fluorouracil combination therapy. *Br J Cancer*, 103(4):581–9.

[140] McLeod, H. L., Sargent, D. J., Marsh, S., Green, E. M., King, C. R., Fuchs, C. S., Ramanathan, R. K., Williamson, S. K., Findlay, B. P., Thibodeau, S. N., Grothey, A., Morton, R. F., and Goldberg, R. M. (2010). Pharmacogenetic predictors of adverse events and response to chemotherapy in metastatic colorectal cancer: results from north american gastrointestinal intergroup trial n9741. *J Clin Oncol*, 28(20):3227–33.

[141] Midgley, R. and Kerr, D. J. (2009). Capecitabine: have we got the dose right? *Nature Clinical Practice Oncology*, 6(1):17–24.

[142] Midgley, R. S., McConkey, C. C., Johnstone, E. C., Dunn, J. A., Smith, J. L., Grumett, S. A., Julier, P., Iveson, C., Yanagisawa, Y., Warren, B., Langman, M. J., and Kerr, D. J. (2010). Phase iii randomized trial assessing rofecoxib in the adjuvant setting of colorectal cancer: final results of the victor trial. *J Clin Oncol*, 28(30):4575–80.

[143] Misset, J. L. (1998). Oxaliplatin in practice. *Br J Cancer*, 77 Suppl 4:4–7.

[144] Miwa, M., Ura, M., Nishida, M., Sawada, N., Ishikawa, T., Mori, K., Shimma, N., Umeda, I., and Ishitsuka, H. (1998). Design of a novel oral fluoropyrimidine carbamate, capecitabine, which generates 5-fluorouracil selectively in tumours by enzymes concentrated in human liver and cancer tissue. *Eur J Cancer*, 34(8):1274–81.

[145] Morel, A., Boisdron-Celle, M., Fey, L., Soulie, P., Craipeau, M. C., Traore, S., and Gamelin, E. (2006). Clinical relevance of different dihydropyrimidine dehydrogenase gene single nucleotide polymorphisms on 5-fluorouracil tolerance. *Mol Cancer Ther*, 5(11):2895–904.

[146] Nica, A. C., Parts, L., Glass, D., Nisbet, J., Barrett, A., Sekowska, M., Travers, M., Potter, S., Grundberg, E., Small, K., Hedman, A. K., Bataille, V., Bell, J. T., Surdulescu, G., Dimas, A. S., Ingle, C., Nestle, F. O., di Meglio, P., Min, J. L., Wilk, A., Hammond, C. J., Hassanali, N., Yang, T. P., Montgomery, S. B., O'Rahilly, S., Lindgren, C. M., Zondervan, K. T., Soranzo, N., Barroso, I., Durbin, R., Ahmadi, K., Deloukas, P., McCarthy, M. I., Dermitzakis, E. T., Spector, T. D., and Consortium, M. (2011). The architecture of gene regulatory variation across multiple human tissues: The muther study. *Plos Genetics*, 7(2).

[147] O'Connell, J. B., Maggard, M. A., and Ko, C. Y. (2004). Colon cancer survival rates with the new american joint committee on cancer sixth edition staging. *J Natl Cancer Inst*, 96(19):1420–5.

[148] O'Connell, M. J. (1989). A phase iii trial of 5-fluorouracil and leucovorin in the treatment of advanced colorectal cancer. a mayo clinic/north central cancer treatment group study. *Cancer*, 63(6 Suppl):1026–30.

[149] Offer, S. M., Wegner, N. J., Fossum, C., Wang, K., and Diasio, R. B. (2013). Phenotypic profiling of dpyd variations relevant to 5-fluorouracil sensitivity using real-time cellular analysis and in vitro measurement of enzyme activity. *Cancer Res.*

[150] Pardini, B., Kumar, R., Naccarati, A., Novotny, J., Prasad, R. B., Forsti, A., Hemminki, K., Vodicka, P., and Lorenzo Bermejo, J. (2011). 5-fluorouracil-based chemotherapy for colorectal cancer and mthfr/mtrr genotypes. *Br J Clin Pharmacol*, 72(1):162–3.

[151] Passarelli, M. N., Coghill, A. E., Hutter, C. M., Zheng, Y., Makar, K. W., Potter, J. D., and Newcomb, P. A. (2011). Common colorectal cancer risk variants in smad7 are associated with survival among prediagnostic nonsteroidal anti-inflammatory drug users: a population-based study of postmenopausal women. *Genes Chromosomes Cancer*, 50(11):875–86.

[152] Peters, G. J., Backus, H. H. J., Freemantle, S., van Triest, B., Codacci-Pisanelli, G., van der Wilt, C. L., Smid, K., Lunec, J., Calvert, A. H., Marsh, S., McLeod, H. L., Bloemena, E., Meijer, S., Jansen, G., van Groeningen, C. J., and Pinedo, H. M. (2002). Induction of thymidylate synthase as a 5-fluorouracil resistance mechanism. *Biochimica Et Biophysica Acta-Molecular Basis of Disease*, 1587(2-3):194–205.

[153] Phipps, A. I., Newcomb, P. A., Garcia-Albeniz, X., Hutter, C. M., White, E., Fuchs, C. S., Hazra, A., Ogino, S., Nan, H., Ma, J., Campbell, P. T., Figueiredo, J. C., Peters, U., and Chan, A. T. (2012). Association between colorectal cancer susceptibility loci and survival time after diagnosis with colorectal cancer. *Gastroenterology*, 143(1):51–4 e4.

[154] Pirinen, M., Donnelly, P., and Spencer, C. C. (2012). Including known covariates can reduce power to detect genetic effects in case-control studies. *Nat Genet*, 44(8):848–51.

[155] Popat, S. and Houlston, R. S. (2005). A systematic review and meta-analysis of

the relationship between chromosome 18q genotype, dcc status and colorectal cancer prognosis. *European Journal of Cancer*, 41(14):2060–2070.

[156] Popat, S., Hubner, R., and Houlston, R. S. (2005). Systematic review of microsatellite instability and colorectal cancer prognosis. *J Clin Oncol*, 23(3):609–18.

[157] Prado, C. M., Baracos, V. E., McCargar, L. J., Mourtzakis, M., Mulder, K. E., Reiman, T., Butts, C. A., Scarfe, A. G., and Sawyer, M. B. (2007). Body composition as an independent determinant of 5-fluorouracil-based chemotherapy toxicity. *Clin Cancer Res*, 13(11):3264–8.

[158] Prahallad, A., Sun, C., Huang, S. D., Di Nicolantonio, F., Salazar, R., Zecchin, D., Beijersbergen, R. L., Bardelli, A., and Bernards, R. (2012). Unresponsiveness of colon cancer to braf(v600e) inhibition through feedback activation of egfr. *Nature*, 483(7387):100–U146.

[159] Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38(8):904–909.

[160] Pullarkat, S. T., Stoehlmacher, J., Ghaderi, V., Xiong, Y. P., Ingles, S. A., Sherrod, A., Warren, R., Tsao-Wei, D., Groshen, S., and Lenz, H. J. (2001). Thymidylate synthase gene polymorphism determines response and toxicity of 5-fu chemotherapy. *Pharmacogenomics J*, 1(1):65–70.

[161] Reigner, B., Watanabe, T., Schuller, J., Lucraft, H., Sasaki, Y., Bridgewater, J., Saeki, T., McAleer, J., Kuranami, M., Poole, C., Kimura, M., Monkhouse, J., Yorulmaz, C., Weidekamm, E., and Grange, S. (2003). Pharmacokinetics of capecitabine (xeloda) in japanese and caucasian patients with breast cancer. *Cancer Chemother Pharmacol*, 52(3):193–201.

[162] Ribelles, N., Lopez-Siles, J., Sanchez, A., Gonzalez, E., Sanchez, M. J., Carabantes, F., Sanchez-Rovira, P., Marquez, A., Duenas, R., Sevilla, I., and Alba, E. (2008). A

carboxylesterase 2 gene polymorphism as predictor of capecitabine on response and time to progression. *Curr Drug Metab*, 9(4):336–43.

[163] Roosild, T. P., Castronovo, S., Villoso, A., Ziemba, A., and Pizzorno, G. (2011). A novel structural mechanism for redox regulation of uridine phosphorylase 2 activity. *J Struct Biol*, 176(2):229–37.

[164] Roth, A. D., Tejpar, S., Delorenzi, M., Yan, P., Fiocca, R., Klingbiel, D., Dietrich, D., Biesmans, B., Bodoky, G., Barone, C., Aranda, E., Nordlinger, B., Cisar, L., Labianca, R., Cunningham, D., Van Cutsem, E., and Bosman, F. (2010). Prognostic role of kras and braf in stage ii and iii resected colon cancer: Results of the translational study on the petacc-3, eortc 40993, sakk 60-00 trial. *Journal of Clinical Oncology*, 28(3):466–474.

[165] Rothenberg, M. L., Meropol, N. J., Poplin, E. A., Van Cutsem, E., and Wadler, S. (2001). Mortality associated with irinotecan plus bolus fluorouracil/leucovorin: summary findings of an independent panel. *J Clin Oncol*, 19(18):3801–7.

[166] Ruzzo, A., Graziano, F., Loupakis, F., Santini, D., Catalano, V., Bisonni, R., Ficarelli, R., Fontana, A., Andreoni, F., Falcone, A., Canestrari, E., Tonini, G., Mari, D., Lippe, P., Pizzagalli, F., Schiavon, G., Alessandroni, P., Giustini, L., Maltese, P., Testa, E., Menichetti, E. T., and Magnani, M. (2008). Pharmacogenetic profiling in patients with advanced colorectal cancer treated with first-line folfiri chemotherapy. *Pharmacogenomics J*, 8(4):278–88.

[167] Saarto, T., Blomqvist, C., Rissanen, P., Auvinen, A., and Elomaa, I. (1997). Haematological toxicity: A marker of adjuvant chemotherapy efficacy in stage ii and iii breast cancer. *British Journal of Cancer*, 75(2):301–305.

[168] Salgado, J., Zabalegui, N., Gil, C., Monreal, I., Rodriguez, J., and Garcia-Foncillas, J. (2007). Polymorphisms in the thymidylate synthase and dihydropyrimidine dehydrogenase genes predict response and toxicity to capecitabine-raltitrexed in colorectal cancer. *Oncol Rep*, 17(2):325–8.

[169] Saltz, L. B., Niedzwiecki, D., Hollis, D., Goldberg, R. M., Hantel, A., Thomas, J. P., Fields, A. L., and Mayer, R. J. (2007). Irinotecan fluorouracil plus leucovorin is not superior to fluorouracil plus leucovorin alone as adjuvant treatment for stage iii colon cancer: results of calgb 89803. *J Clin Oncol*, 25(23):3456–61.

[170] Sanger, F., Nicklen, S., and Coulson, A. R. (1977). Dna sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, 74(12):5463–7.

[171] Sargent, D., Sobrero, A., Grothey, A., O'Connell, M. J., Buyse, M., Andre, T., Zheng, Y., Green, E., Labianca, R., O'Callaghan, C., Seitz, J. F., Francini, G., Haller, D., Yothers, G., Goldberg, R., and de Gramont, A. (2009). Evidence for cure by adjuvant therapy in colon cancer: observations based on individual patient data from 20,898 patients on 18 randomized trials. *J Clin Oncol*, 27(6):872–7.

[172] Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression-model. *Biometrika*, 69(1):239–241.

[173] Schwab, M., Zanger, U. M., Marx, C., Schaeffeler, E., Klein, K., Dippon, J., Kerb, R., Blievernicht, J., Fischer, J., Hofmann, U., Bokemeyer, C., and Eichelbaum, M. (2008). Role of genetic and nongenetic factors for fluorouracil treatment-related severe toxicity: a prospective clinical trial by the german 5-fu toxicity study group. *J Clin Oncol*, 26(13):2131–8.

[174] Seck, K., Riemer, S., Kates, R., Ullrich, T., Lutz, V., Harbeck, N., Schmitt, M., Kiechle, M., Diasio, R., and Gross, E. (2005). Analysis of the dpyd gene implicated in 5-fluorouracil catabolism in a cohort of caucasian individuals. *Clin Cancer Res*, 11(16):5886–92.

[175] Sharma, R., Hoskins, J. M., Rivory, L. P., Zucknick, M., London, R., Liddle, C., and Clarke, S. J. (2008). Thymidylate synthase and methylenetetrahydrofolate reductase gene polymorphisms and toxicity to capecitabine in advanced colorectal cancer patients. *Clin Cancer Res*, 14(3):817–25.

[176] Sharma, R., Rivory, L., Beale, P., Ong, S., Horvath, L., and Clarke, S. J. (2006). A phase ii study of fixed-dose capecitabine and assessment of predictors of toxicity in patients with advanced/metastatic colorectal cancer. *Br J Cancer*, 94(7):964–8.

[177] Shu, X. O., Long, J., Lu, W., Li, C., Chen, W. Y., Delahanty, R., Cheng, J., Cai, H., Zheng, Y., Shi, J., Gu, K., Wang, W. J., Kraft, P., Gao, Y. T., Cai, Q., and Zheng, W. (2012). Novel genetic markers of breast cancer survival identified by a genome-wide association study. *Cancer Res*, 72(5):1182–9.

[178] Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 13(2):238–241.

[179] Singer, V. L., Jones, L. J., Yue, S. T., and Haugland, R. P. (1997). Characterization of picogreen reagent and development of a fluorescence-based solution assay for double-stranded dna quantitation. *Anal Biochem*, 249(2):228–38.

[180] Sklar, P., Gabriel, S. B., McInnis, M. G., Bennett, P., Lim, Y. M., Tsan, G., Schaffner, S., Kirov, G., Jones, I., Owen, M., Craddock, N., DePaulo, J. R., and Lander, E. S. (2002). Family-based association study of 76 candidate genes in bipolar disorder: Bdnf is a potential risk locus. *Molecular Psychiatry*, 7(6):579–593.

[181] Sloan, J. A., Goldberg, R. M., Sargent, D. J., Vargas-Chanes, D., Nair, S., Cha, S. S., Novotny, P. J., Poon, M. A., O'Connell, M. J., and Loprinzi, C. L. (2002). Women experience greater toxicity with fluorouracil-based chemotherapy for colorectal cancer. *J Clin Oncol*, 20(6):1491–8.

[182] So, H. C., Gui, A. H. S., Cherny, S. S., and Sham, P. C. (2011). Evaluating the heritability explained by known susceptibility variants: A survey of ten complex diseases. *Genetic Epidemiology*, 35(5):310–317.

[183] Speed, D., Hemani, G., Johnson, M. R., and Balding, D. J. (2012). Improved heritability estimation from genome-wide snps. *American Journal of Human Genetics*, 91(6):1011–1021.

[184] Spurdle, A. B., Thompson, D. J., Ahmed, S., Ferguson, K., Healey, C. S., O'Mara, T., Walker, L. C., Montgomery, S. B., Dermitzakis, E. T., Fahey, P., Montgomery, G. W., Webb, P. M., Fasching, P. A., Beckmann, M. W., Ekici, A. B., Hein, A., Lambrechts, D., Coenegrachts, L., Vergote, I., Amant, F., Salvesen, H. B., Trovik, J., Njolstad, T. S., Helland, H., Scott, R. J., Ashton, K., Proietto, T., Otton, G., Tomlinson, I., Gorman, M., Howarth, K., Hodgson, S., Garcia-Closas, M., Wentzensen, N., Yang, H. N., Chanock, S., Hall, P., Czene, K., Liu, J. J., Li, J. M., Shu, X. O., Zheng, W., Long, J. R., Xiang, Y. B., Shah, M., Morrison, J., Michailidou, K., Pharoah, P. D., Dunning, A. M., Easton, D. F., Canc, A. N. E., and Geneti, N. S. E. C. (2011). Genome-wide association study identifies a common variant associated with risk of endometrial cancer. *Nature Genetics*, 43(5):451–+.

[185] Srinivasan, Y., Sasa, M., Honda, J., Takahashi, A., Uno, S., Kamatani, N., Kubo, M., Nakamura, Y., and Zembutsu, H. (2011). Genome-wide association study of epirubicin-induced leukopenia in japanese patients. *Pharmacogenetics and Genomics*, 21(9):552–558.

[186] Stein, B. N., Petrelli, N. J., Douglass, H. O., Driscoll, D. L., Arcangeli, G., and Meropol, N. J. (1995). Age and sex are independent predictors of 5-fluorouracil toxicity. analysis of a large scale phase iii trial. *Cancer*, 75(1):11–7.

[187] Sterne, J. A., Gavaghan, D., and Egger, M. (2000). Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol*, 53(11):1119–29.

[188] Stranger, B. E., Montgomery, S. B., Dimas, A. S., Parts, L., Stegle, O., Ingle, C. E., Sekowska, M., Smith, G. D., Evans, D., Gutierrez-Arcelus, M., Price, A., Raj, T., Nisbett, J., Nica, A. C., Beazley, C., Durbin, R., Deloukas, P., and Dermitzakis, E. T. (2012). Patterns of cis regulatory variation in diverse human populations. *Plos Genetics*, 8(4):272–284.

[189] Stranger, B. E., Stahl, E. A., and Raj, T. (2011). Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics*, 187(2):367–83.

[190] Tabor, H. K., Risch, N. J., and Myers, R. M. (2002). Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nature Reviews Genetics*, 3(5):391–A396.

[191] Taniuchi, I., Osato, M., and Ito, Y. (2012). Runx1: no longer just for leukemia. *Embo Journal*, 31(21):4098–4099.

[192] Tenesa, A., Theodoratou, E., Din, F. V., Farrington, S. M., Cetnarskyj, R., Barnetson, R. A., Porteous, M. E., Campbell, H., and Dunlop, M. G. (2010). Ten common genetic variants associated with colorectal cancer risk are not associated with survival after diagnosis. *Clin Cancer Res*, 16(14):3754–9.

[193] Thompson, S. G. and Pocock, S. J. (1991). Can meta-analyses be trusted? *Lancet*, 338(8775):1127–30.

[194] Thorn, C. F., Marsh, S., Carrillo, M. W., McLeod, H. L., Klein, T. E., and Altman, R. B. (2011). Pharmgkb summary: fluoropyrimidine pathways. *Pharmacogenet Genomics*, 21(4):237–42.

[195] Tomlinson, I., Webb, E., Carvajal-Carmona, L., Broderick, P., Kemp, Z., Spain, S., Penegar, S., Chandler, I., Gorman, M., Wood, W., Barclay, E., Lubbe, S., Martin, L., Sellick, G., Jaeger, E., Hubner, R., Wild, R., Rowan, A., Fielding, S., Howarth, K., Silver, A., Atkin, W., Muir, K., Logan, R., Kerr, D., Johnstone, E., Sieber, O., Gray, R., Thomas, H., Peto, J., Cazier, B., Houlston, R., and Consortium, C. (2007). A genome-wide association scan of tag snps identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nature Genetics*, 39(8):984–988.

[196] Tomlinson, I. P., Carvajal-Carmona, L. G., Dobbins, S. E., Tenesa, A., Jones, A. M., Howarth, K., Palles, C., Broderick, P., Jaeger, E. E., Farrington, S., Lewis, A., Prendergast, J. G., Pittman, A. M., Theodoratou, E., Olver, B., Walker, M., Penegar, S., Barclay, E., Whiffin, N., Martin, L., Ballereau, S., Lloyd, A., Gorman, M., Lubbe, S., Consortium, C., Collaborators, C., Consortium, E., Howie, B., Marchini, J., Ruiz-Ponte,

C., Fernandez-Rozadilla, C., Castells, A., Carracedo, A., Castellvi-Bel, S., Duggan, D., Conti, D., Cazier, J. B., Campbell, H., Sieber, O., Lipton, L., Gibbs, P., Martin, N. G., Montgomery, G. W., Young, J., Baird, P. N., Gallinger, S., Newcomb, P., Hopper, J., Jenkins, M. A., Aaltonen, L. A., Kerr, D. J., Cheadle, J., Pharoah, P., Casey, G., Houlston, R. S., and Dunlop, M. G. (2011). Multiple common susceptibility variants near bmp pathway loci grem1, bmp4, and bmp2 explain part of the missing heritability of colorectal cancer. *PLoS Genet*, 7(6):e1002105.

[197] Tomlinson, I. P., Webb, E., Carvajal-Carmona, L., Broderick, P., Howarth, K., Pittman, A. M., Spain, S., Lubbe, S., Walther, A., Sullivan, K., Jaeger, E., Fielding, S., Rowan, A., Vijayakrishnan, J., Domingo, E., Chandler, I., Kemp, Z., Qureshi, M., Farrington, S. M., Tenesa, A., Prendergast, J. G., Barnetson, R. A., Penegar, S., Barclay, E., Wood, W., Martin, L., Gorman, M., Thomas, H., Peto, J., Bishop, D. T., Gray, R., Maher, E. R., Lucassen, A., Kerr, D., Evans, D. G., Consortium, C., Schafmayer, C., Buch, S., Volzke, H., Hampe, J., Schreiber, S., John, U., Koessler, T., Pharoah, P., van Wezel, T., Morreau, H., Wijnen, J. T., Hopper, J. L., Southey, M. C., Giles, G. G., Severi, G., Castellvi-Bel, S., Ruiz-Ponte, C., Carracedo, A., Castells, A., Consortium, E., Forsti, A., Hemminki, K., Vodicka, P., Naccarati, A., Lipton, L., Ho, J. W., Cheng, K. K., Sham, P. C., Luk, J., Agundez, J. A., Ladero, J. M., de la Hoya, M., Caldes, T., Niittymaki, I., Tuupanen, S., Karhu, A., Aaltonen, L., Cazier, J. B., Campbell, H., Dunlop, M. G., and Houlston, R. S. (2008). A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet*, 40(5):623–30.

[198] Trevino, L. R., Shimasaki, N., Yang, W., Panetta, J. C., Cheng, C., Pei, D., Chan, D., Sparreboom, A., Giacomini, K. M., Pui, C. H., Evans, W. E., and Relling, M. V. (2009). Germline genetic variation in an organic anion transporter polypeptide associated with methotrexate pharmacokinetics and clinical effects. *J Clin Oncol*, 27(35):5972–8.

[199] Twelves, C., Wong, A., Nowacki, M. P., Abt, M., Burris, H., r., Carrato, A., Cassidy, J., Cervantes, A., Fagerberg, J., Georgoulias, V., Husseini, F., Jodrell, D., Koralewski, P., Kroning, H., Maroun, J., Marschner, N., McKendrick, J., Pawlicki, M., Rosso, R.,

Schuller, J., Seitz, J. F., Stabuc, B., Tujakowski, J., Van Hazel, G., Zaluski, J., and Scheithauer, W. (2005). Capecitabine as adjuvant treatment for stage iii colon cancer. *N Engl J Med*, 352(26):2696–704.

[200] Twelves, C. J. (2006). Xeloda in adjuvant colon cancer therapy (x-act) trial: overview of efficacy, safety, and cost-effectiveness. *Clin Colorectal Cancer*, 6(4):278–87.

[201] Van Cutsem, E., Labianca, R., Bodoky, G., Barone, C., Aranda, E., Nordlinger, B., Topham, C., Tabernero, J., Andre, T., Sobrero, A. F., Mini, E., Greil, R., Di Costanzo, F., Collette, L., Cisar, L., Zhang, X., Khayat, D., Bokemeyer, C., Roth, A. D., and Cunningham, D. (2009). Randomized phase iii trial comparing biweekly infusional fluorouracil/leucovorin alone or with irinotecan in the adjuvant treatment of stage iii colon cancer: Petacc-3. *J Clin Oncol*, 27(19):3117–25.

[202] Van Cutsem, E., Twelves, C., Cassidy, J., Allman, D., Bajetta, E., Boyer, M., Bugat, R., Findlay, M., Frings, S., Jahn, M., McKendrick, J., Osterwalder, B., Perez-Manga, G., Rosso, R., Rougier, P., Schmiegel, W. H., Seitz, J. F., Thompson, P., Vieitez, J. M., Weitzel, C., Harper, P., and Xeloda Colorectal Cancer Study, G. (2001). Oral capecitabine compared with intravenous fluorouracil plus leucovorin in patients with metastatic colorectal cancer: results of a large phase iii study. *J Clin Oncol*, 19(21):4097–106.

[203] van Kuilenburg, A. B., Dobritzsch, D., Meinsma, R., Haasjes, J., Waterham, H. R., Nowaczyk, M. J., Maropoulos, G. D., Hein, G., Kalhoff, H., Kirk, J. M., Baaske, H., Aukett, A., Duley, J. A., Ward, K. P., Lindqvist, Y., and van Gennip, A. H. (2002). Novel disease-causing mutations in the dihydropyrimidine dehydrogenase gene interpreted by analysis of the three-dimensional protein structure. *Biochem J*, 364(Pt 1):157–63.

[204] van Kuilenburg, A. B., Haasjes, J., Richel, D. J., Zoetekouw, L., Van Lenthe, H., De Abreu, R. A., Maring, J. G., Vreken, P., and van Gennip, A. H. (2000). Clinical implications of dihydropyrimidine dehydrogenase (dpd) deficiency in patients with severe

5-fluorouracil-associated toxicity: identification of new mutations in the dpd gene. *Clin Cancer Res*, 6(12):4705–12.

[205] Van Kuilenburg, A. B., Meinsma, R., Beke, E., Bobba, B., Boffi, P., Enns, G. M., Witt, D. R., and Dobritzsch, D. (2005). Identification of three novel mutations in the dihydropyrimidine dehydrogenase gene associated with altered pre-mrna splicing or protein function. *Biol Chem*, 386(4):319–24.

[206] Van Kuilenburg, A. B., Vreken, P., Abeling, N. G., Bakker, H. D., Meinsma, R., Van Lenthe, H., De Abreu, R. A., Smeitink, J. A., Kayserili, H., Apak, M. Y., Christensen, E., Holopainen, I., Pulkki, K., Riva, D., Botteon, G., Holme, E., Tulinius, M., Kleijer, W. J., Beemer, F. A., Duran, M., Niezen-Koning, K. E., Smit, G. P., Jakobs, C., Smit, L. M., Van Gennip, A. H., and et al. (1999). Genotype and phenotype in patients with dihydropyrimidine dehydrogenase deficiency. *Hum Genet*, 104(1):1–9.

[207] Vreken, P., Van Kuilenburg, A. B., Meinsma, R., and van Gennip, A. H. (1997). Dihydropyrimidine dehydrogenase (dpd) deficiency: identification and expression of missense mutations c29r, r886h and r235w. *Hum Genet*, 101(3):333–8.

[208] Walther, A., Houlston, R., and Tomlinson, I. (2008). Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis. *Gut*, 57(7):941–950.

[209] Weisberg, I., Tran, P., Christensen, B., Sibani, S., and Rozen, R. (1998). A second genetic polymorphism in methylenetetrahydrofolate reductase (mthfr) associated with decreased enzyme activity. *Mol Genet Metab*, 64(3):169–72.

[210] Wellcome Trust Case Control, C. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, 447(7145):661–78.

[211] West, C. M., Jones, T., and Price, P. (2004). The potential of positron-emission tomography to study anticancer-drug resistance. *Nat Rev Cancer*, 4(6):457–69.

[212] Whitman, S. P., Maharry, K., Radmacher, M. D., Becker, H., Mrozek, K., Margeson, D., Holland, K. B., Wu, Y. Z., Schwind, S., Metzeler, K. H., Wen, J., Baer, M. R., Powell,

B. L., Carter, T. H., Kolitz, J. E., Wetzler, M., Moore, J. O., Stone, R. M., Carroll, A. J., Larson, R. A., Caligiuri, M. A., Marcucci, G., and Bloomfield, C. D. (2010). Flt3 internal tandem duplication associates with adverse outcome and gene-and microrna-expression signatures in patients 60 years of age or older with primary cytogenetically normal acute myeloid leukemia: a cancer and leukemia group b study. *Blood*, 116(18):3622–3626.

[213] Wolmark, N., Rockette, H., Fisher, B., Wickerham, D. L., Redmond, C., Fisher, E. R., Jones, J., Mamounas, E. P., Ore, L., Petrelli, N. J., and et al. (1993). The benefit of leucovorin-modulated fluorouracil as postoperative adjuvant therapy for primary colon cancer: results from national surgical adjuvant breast and bowel project protocol c-03. *J Clin Oncol*, 11(10):1879–87.

[214] Wray, N. R., Yang, J., Goddard, M. E., and Visscher, P. M. (2010). The genetic interpretation of area under the roc curve in genomic profiling. *PLoS Genet*, 6(2):e1000864.

[215] Wu, M. H., Chen, P., Wu, X., Liu, W., Strom, S., Das, S., Cook, E. H., J., Rosner, G. L., and Dolan, M. E. (2004). Determination and analysis of single nucleotide polymorphisms and haplotype structure of the human carboxylesterase 2 gene. *Pharmacogenetics*, 14(9):595–605.

[216] Xing, J., Myers, R. E., He, X., Qu, F., Zhou, F., Ma, X., Hyslop, T., Bao, G., Wan, S., Yang, H., and Chen, Z. (2011). Gwas-identified colorectal cancer susceptibility locus associates with disease prognosis. *Eur J Cancer*, 47(11):1699–707.

[217] Yamauchi, M., Morikawa, T., Kuchiba, A., Imamura, Y., Qian, Z. R., Nishihara, R., Liao, X. Y., Waldron, L., Hoshida, Y., Huttenhower, C., Chan, A. T., Giovannucci, E., Fuchs, C., and Ogino, S. (2012). Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. *Gut*, 61(6):847–854.

[218] Yang, J. A., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). Gcta: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, 88(1):76–82.

[219] Zalcberg, J., Kerr, D., Seymour, L., and Palmer, M. (1998). Haematological and non-haematological toxicity after 5-fluorouracil and leucovorin in patients with advanced colorectal cancer is significantly associated with gender, increasing age and cycle number. tomudex international study group. *Eur J Cancer*, 34(12):1871–5.

[220] Zarate, R., Rodriguez, J., Bandres, E., Patino-Garcia, A., Ponz-Sarvise, M., Viudez, A., Ramirez, N., Bitarte, N., Chopitea, A., and Gacia-Foncillas, J. (2010). Oxaliplatin, irinotecan and capecitabine as first-line therapy in metastatic colorectal cancer (mcrc): a dose-finding study and pharmacogenomic analysis. *Br J Cancer*, 102(6):987–94.

[221] Zhang, H. T., Xu, Y., Zhang, Z. H., Liu, R. L., and Ma, B. J. (2012). Association between cox-2 rs2745557 polymorphism and prostate cancer risk: a systematic review and meta-analysis. *Bmc Immunology*, 13.

[222] Zhang, Q., Zhao, Y. P., Liao, Q., Hu, Y., Xu, Q., Zhou, L., and Shu, H. (2011). Associations between gene polymorphisms of thymidylate synthase with its protein expression and chemosensitivity to 5-fluorouracil in pancreatic carcinoma cells. *Chin Med J (Engl)*, 124(2):262–7.

[223] Zhang, W., Gordon, M., Schultheis, A. M., Yang, D. Y., Nagashima, F., Azuma, M., Chang, H. M., Borucka, E., Lurje, G., Sherrod, A. E., Iqbal, S., Groshen, S., and Lenz, H. J. (2007). Fcgr2a and fcgr3a polymorphisms associated with clinical outcome of epidermal growth factor receptor expressing metastatic colorectal cancer patients treated with single-agent cetuximab. *J Clin Oncol*, 25(24):3712–8.

[224] Zintzaras, E., Ziogas, D. C., Kitsios, G. D., Papathanasiou, A. A., Lau, J., and Raman, G. (2009). Mthfr gene polymorphisms and response to chemotherapy in colorectal cancer: a meta-analysis. *Pharmacogenomics*, 10(8):1285–94.