Internet use and health:
a mixed methods analysis using spatial microsimulation and interviews

Thesis submitted in partial fulfilment of the requirements for the degree of DPhil in Information, Communication and the Social Sciences in the Oxford Internet Institute at the University of Oxford

Ulrike Deetjen

Balliol College



99,410 words

Supervisors: Dr Rebecca Eynon & Professor John Powell

Trinity 2016

Abstract

Internet use is considered a lever for empowering patients, levelling inequalities and reducing healthcare expenditure. However, with digital inclusion, health provision quality and health system efficiency high on the UK and EU policy agendas, we need to better understand the relationship between Internet use and health outcomes to assess potential benefits and adverse effects.

This research addresses the question of how Internet use influences individuals' health service use and their perceived health in the context of England. Focusing on health information-seeking, it analyses variations across different kinds of users, mechanisms between Internet use and both health outcomes, and the role of individual and contextual factors in this relationship.

To answer this question, this research uses a mixed methods approach. Quantitative data from the Oxford Internet Surveys (OxIS), the English census and Hospital Episode Statistics (HES) was connected through spatial microsimulation based on output areas. Qualitative data was collected through semi-structured, face-to-face interviews, primarily with former OxIS participants from output areas in the quantitative strand. The quantitative data was revisited based on emerging interview themes.

The results indicate that Internet use influences perceived health and health service use via various mechanisms based on the Internet's content, mediation and connection affordances. However, the boundaries between users and non-users are blurry and outcomes vary for different types of individuals, classified here as learners, pragmatists, sceptics, worriers, delegators and adigitals. Age, education, socioeconomic status, long-term health conditions, and geographic context influence Internet use and health outcomes separately, while the social context shapes their relationship too.

These findings advance the theoretical understanding of Internet-based health outcomes, and provide practical implications for health professionals and policy-makers with insights down to the local level. Moreover, this research demonstrates how novel insights for public wellbeing can be derived from qualitatively enriched secondary data in privacy-preserving and cost-effective ways.

Acknowledgements

First and foremost, I wish to thank my supervisors, Dr Rebecca Eynon and Professor John Powell, for being trusted mentors over the last three years. Their support was not only instrumental in shaping this thesis and steering it smoothly through the challenges, but also in terms of my learning the mechanics of academia. Above all, they made the DPhil a very rewarding experience – thank you.

My sincerest gratitude also goes to my examiners, Professor Sue Ziebland and Dr Ben Anderson, as well as to Dr Grant Blank, who assessed the research proposal of this thesis. As my thesis benefitted greatly from their work throughout, I feel privileged that they were the ones examining it.

I am also grateful to my funders: the Clarendon Fund, the Economic and Social Research Council (ESRC), and Balliol College, as well as Dr Victoria Nash who played a crucial role in obtaining this funding. I also thank McKinsey & Company for being supportive of my plan to combine my work with academic research, and for granting me a second period of educational leave to pursue the doctorate.

This research would have been impossible without the 43 anonymous interviewees who generously devoted their time and openly disclosed intimate aspects of their lives online and offline. Their stories were key to giving this thesis a human face.

Moreover, I would like to thank those who shaped particular aspects of the thesis. Professor Danny Dorling and Dr Dimitris Ballas constructively challenged my research design, Dr Robin Lovelace taught me the basics of spatial microsimulation and provided his R code, Dr Lars-Erik Malmberg offered guidance on the structural equation models, and Dr Chrysanthi Papoutsis provided input for the analytical framework. For their comments on the policy recommendations, I thank Dr Pēteris Zilgalvis (European Commission), Bob Gann and Helen Rowntree (NHS England), Mahesh Patel (Public Health England) and Dr Kate Mason (Tinder Foundation).

A few people also deserve a big thank you for their help with accessing Hospital Episode Statistics: Professor Martin Landray and Sarah Lawson for their guidance, Adham Tamer for writing various new security policies and setting up the technical infrastructure, and Eve Henshaw for bringing order to the contractual chaos.

Finally, I would like to thank the directors of the DPhil programme, Dr Eric Meyer and Dr Vili Lehdonvirta, and in particular my DPhil colleagues, who provided valuable feedback in our weekly seminars and informal conversations. I also thank everyone in the Oxford Internet Institute and the Balliol College community, who made my time in Oxford an incredible experience.

Last but not least, I wish to thank my family – in particular my parents, my sister Luise, my grandparents, my parents and sister-in-law, and most importantly, my husband Jochen – for their love and constant support over the years.

Table of Contents

List of Figures **6**

List of Tables **7**

List of Abbreviations **8**

1 Introduction **11**
- 1.1 Overview and research questions . . . 13
- 1.2 Background . . . 15
  - 1.2.1 Internet use in England . . . 15
  - 1.2.2 Healthcare in England . . . 16
- 1.3 Contributions of this research . . . 19
  - 1.3.1 Theoretical contribution . . . 19
  - 1.3.2 Practical contribution . . . 20
  - 1.3.3 Methodological contribution . . . 21
- 1.4 Thesis structure . . . 22

2 Literature review **25**
- 2.1 Health . . . 28
  - 2.1.1 Models of health service use . . . 30
  - 2.1.2 Models of perceived health . . . 33
- 2.2 Internet use . . . 36
  - 2.2.1 Three user concepts based on the digital divides . . . 37
  - 2.2.2 The importance of antecedent factors . . . 39
  - 2.2.3 Motivations for health-related Internet use . . . 41
  - 2.2.4 Enablers for health-related Internet use . . . 43
  - 2.2.5 RQ_1: Health information for (non-)users . . . 43
- 2.3 Effects of Internet use on health . . . 44
  - 2.3.1 Affordances: Content, mediation and connection . . . 45
  - 2.3.2 Resources: Information and knowledge capital . . . 47
  - 2.3.3 Resources: Social support and social capital . . . 50
  - 2.3.4 RQ_2: Mechanisms of Internet use and health . . . 52

2.4 Individual and contextual factors . . . 52
2.4.1 Demographics . . . 54
2.4.2 Social environment . . . 58
2.4.3 Geographic environment . . . 59
2.4.4 RQ_3: Individual and contextual influences . . . 61

3 Methodology 63
3.1 Mixed methods research design . . . 64
3.2 Quantitative data formation . . . 66
3.2.1 Selection of datasets . . . 67
3.2.2 Combination through spatial microsimulation . . . 70
3.2.3 Internal validation . . . 76
3.2.4 External validation . . . 78
3.3 Qualitative data collection . . . 81
3.3.1 Sampling approach . . . 81
3.3.2 Interview design . . . 86
3.4 Analytical strategy . . . 89
3.4.1 Interpretation of interview insights . . . 90
3.4.2 Development of typologies . . . 91
3.4.3 Analysis of large-scale simulated data . . . 93
3.4.4 Validity and reliability . . . 96
3.5 Ethical considerations . . . 100

4 Results: Health information for (non-)users 103
4.1 Overview of perceived health and health service use . . . 105
4.2 Three user concepts and health . . . 107
4.2.1 Access divide . . . 107
4.2.2 Support divide . . . 109
4.2.3 Usage divide . . . 112
4.3 Antecedent factors for health-related Internet use . . . 115
4.3.1 Motivations . . . 115
4.3.2 Enablers . . . 120
4.3.3 Bridging qualitative and quantitative data . . . 123
4.4 A typology of health information-seekers . . . 125
4.4.1 An overview of the six types . . . 127
4.4.2 The learners . . . 131
4.4.3 The pragmatists . . . 133
4.4.4 The sceptics . . . 136
4.4.5 The worriers . . . 137
4.4.6 The delegators . . . 140

4.4.7 The adigitals . . . 142
4.5 Effects of Internet use across types . . . 143
4.5.1 Perceived health . . . 145
4.5.2 Health service use . . . 148

5 Results: Mechanisms of Internet use and health 151
5.1 Effects on health service use . . . 153
5.1.1 Avoiding medical help . . . 154
5.1.2 Encouraging service use . . . 156
5.1.3 Estimating urgency . . . 157
5.1.4 Monitoring for emergencies . . . 158
5.1.5 Spurring discussions in consultations . . . 160
5.1.6 Being more demanding in consultations . . . 162
5.1.7 Making decisions about treatments . . . 162
5.1.8 Affecting health service satisfaction . . . 164
5.2 Effects on perceived health . . . 165
5.2.1 Re-evaluating a situation based on information . . . 165
5.2.2 Comparing and feeling normal . . . 168
5.2.3 Exchanging best practices with others . . . 170
5.2.4 Affecting perception of lifestyles . . . 172
5.2.5 Managing identity and boundaries . . . 173
5.2.6 Finding companionship . . . 175
5.2.7 Being embedded with family and friends . . . 178
5.2.8 Changing life circumstances . . . 179
5.3 Quantitative models of selected mechanisms . . . 180
5.3.1 Health outcomes of general usage . . . 181
5.3.2 Health outcomes of social connections . . . 184
5.3.3 Health outcomes of health information-seeking . . . 188

6 Results: Individual and contextual influences 191
6.1 Demographics . . . 192
6.1.1 Age . . . 193
6.1.2 Education . . . 196
6.1.3 Socioeconomic status . . . 200
6.1.4 Long-term health conditions . . . 203
6.2 Social environment . . . 206
6.2.1 Compensating for a lack of enablers . . . 207
6.2.2 Filtering information . . . 208
6.3 Geographic environment . . . 211
6.3.1 Area-based differences . . . 212

6.3.2 Country-based differences . . . 216
6.3.3 Case study: Isle of Wight . . . 218

7 Discussion **223**
7.1 Multiple mechanisms between the Internet and health . . . 225
7.1.1 Affordances . . . 226
7.1.2 Resources . . . 228
7.1.3 Health factors . . . 232
7.2 Delineating the shifting digital divide . . . 237
7.2.1 Towards an outcomes divide . . . 237
7.2.2 Social and digital inclusion . . . 240
7.2.3 Non-users' mechanisms . . . 242
7.2.4 A new lifestyle paradox? . . . 244
7.3 The duality of agency and structure . . . 246
7.3.1 Creating and managing fields . . . 247
7.3.2 Levelling the field in the consultation . . . 249
7.3.3 Shaping sociodemographic structure . . . 251
7.3.4 Reconfiguring spatial structures . . . 252
7.4 Implications for practice and policy . . . 253
7.4.1 Enabling outcomes while leaving nobody behind . . . 254
7.4.2 Shifting responsibility from the state to the individual . . . 257
7.4.3 Allowing data-driven insights for health and healthcare . . . 259
7.5 Methods of enquiry within the pragmatic paradigm . . . 261
7.5.1 Mixed methods research design . . . 261
7.5.2 Obtaining qualitative health data . . . 263
7.5.3 Spatial microsimulation for combining datasets . . . 265
7.5.4 Recommendations for secondary data . . . 268

8 Conclusion **271**
8.1 Synthesis of research questions . . . 272
8.1.1 RQ_1: Health information for (non-)users . . . 272
8.1.2 RQ_2: Mechanisms of Internet use and health . . . 274
8.1.3 RQ_3: Individual and contextual influences . . . 276
8.2 Contribution of findings . . . 277
8.2.1 Theoretical contributions . . . 277
8.2.2 Practical contributions . . . 279
8.2.3 Methodological contributions . . . 280
8.3 Directions for further investigation . . . 282

Appendices **285**

Literature review **287**

I Overview of previous studies . . . 287
II Key elements of theories on health service use . . . 291
III Key elements of theories on health perception . . . 294

Methodology (quantitative) **295**

IV Mapping of theoretical constructs to datasets . . . 295
V Constraint variables for spatial microsimulation . . . 302
VI Construction of outcome variables . . . 306
VII Internal validation results . . . 308
VIII External validation results . . . 310
IX Spatial microsimulation in R: approach and code . . . 316
X Regression diagnostics . . . 328
XI Results for OxIS versus simulated dataset . . . 332
XII Timeline of HES data request . . . 339

Methodology (qualitative) **340**

XIII Area selection . . . 340
XIV Interviewee health conditions . . . 341
XV Overview of interviewees . . . 342
XVI Interview participant documents . . . 346
XVII Interview guide . . . 349
XVIII Coding scheme . . . 352
XIX Research summary and interviewee responses . . . 357

Dissemination **361**

XX Policy summary . . . 361
XXI Overview of publications . . . 363

Bibliography **365**

List of Figures

Figure 1: Overview of analytical framework 27
Figure 2: Analytical framework (1/4) – influences on health 30
Figure 3: Analytical framework (2/4) – enablers and motivations 40
Figure 4: Analytical framework (3/4) – affordances and resources 45
Figure 5: Analytical framework (4/4) – individual/contextual factors . . 53
Figure 6: Mapping of research questions to analytical framework 62
Figure 7: Schematic summary of spatial microsimulation 74
Figure 8: Overview of selected areas 83
Figure 9: SEM based on access divide 108
Figure 10: SEM based on support divide 110
Figure 11: SEM based on usage divide 112
Figure 12: Dendrogram for hierarchical clustering of typology 125
Figure 13: Typology of health information seekers – cluster dimensions . 126
Figure 14: SEM based on usage divide – split up by typology 145
Figure 15: Effects on health service use based on qualitative research . . 154
Figure 16: Effects on perceived health based on qualitative research . . . 165
Figure 17: SEM for "affecting life circumstances" mechanism 181
Figure 18: Multinomial regression for "affecting lifestyle" mechanism . . 183
Figure 19: SEM for "finding companionship" mechanism 184
Figure 20: SEM for "being embedded with family/friends" mechanism . . 186
Figure 21: SEM for "monitoring for emergencies" mechanism 187
Figure 22: SEM for health information seeking mechanisms 189
Figure 23: Maps of outcome concepts per local authority district 211
Figure 24: Maps of Internet-based effects regressed on perceived health . 215
Figure 25: Maps of Internet use and health for OAs on the Isle of Wight 220
Figure 26: Theoretical affordances versus observed mechanisms 226
Figure 27: Theoretical resources versus observed mechanisms 229
Figure 28: Theoretical health factors versus observed mechanisms 233
Figure 29: Conceptual overview of lifestyle paradox 244
Figure 30: Updated mediating mechanisms in analytical framework . . . 275

List of Tables

Table 1: User concepts following three digital divides 37
Table 2: Affordances of Internet use 47
Table 3: Mixed methods research design 64
Table 4: Employed datasets matched to analytical framework 68
Table 5: Simulated vs. actual values for South East region 76
Table 6: External validation – comparison of proportions of individuals . 79
Table 7: External validation – average perceived health status 80
Table 8: Area selection logic 82
Table 9: Interviewee demographic characteristics 85
Table 10: Descriptive statistics for outcome variables 105
Table 11: Regression based on access divide 107
Table 12: Regression based on support divide 110
Table 13: Typology of health information seekers – other characteristics . 129
Table 14: Logistic regression for types of health information seekers . . . 130
Table 15: Regression with typology of health information seekers 144
Table 16: Regression for age effects 195
Table 17: Regression for education effects 199
Table 18: Regression for socio-economic status effects 202
Table 19: Regression for long-term health condition effects 205
Table 20: Regression for social environment effects 210
Table 21: Regression on perceived health across regions 212
Table 22: Regression on perceived health across urban/rural areas 213
Table 23: Regression on health service use across urban/rural areas . . . 214
Table 24: Regression for access, support and usage divide (Isle of Wight) 219
Table 25: Practical implications based on typology 255

List of Abbreviations

ADHD	Attention Deficit Hyperactivity Disorder
AMU	Adult Media Use survey
BPPV	Benign Paroxysmal Positional Vertigo
BHPS	British Household Panel Survey
BSA	British Sociological Association
CFI	Comparative Fit Index
CI	Condition Index
CPRD	Clinical Practice Research Database
COPD	Chronic Obstructive Pulmonary Disease
DAAG	Data Access Approval Group
DVT	Deep Vein Thrombosis
ESRC	Economic and Social Research Council
ESS	Error of the Sum of Squares
EU	European Union
FOIA	Freedom of Information Act
GBP	Great British Pound
GP	General Practitioner
HES	Hospital Episode Statistics
HSCIC	Health & Social Care Information Centre
IBS	Irritable Bowel Syndrome
ICD	International Classification of Diseases
IM	Instant Messaging
IP	Internet Protocol
IPF	Iterative Proportional Fitting
IVF	In-vitro Fertilisation
LOWESS	Locally Weighted Scatterplot Smoothing

MAUP	Modifiable Areal Unit Problem
MCD	Major and Chronic Diseases
MRC	Medical Research Council
MS	Multiple Sclerosis
NHS	National Health Service
NIB	National Information Board
NRMSE	Normalised Root Mean Standardised Error
NS-SeC	National Statistics Socioeconomic Classification
NVQ	National Vocational Qualification
OA	Output Area
OCD	Obsessive Compulsory Disease
OII	Oxford Internet Institute
OLS	Opinions & Lifestyle Survey
ONS	Office for National Statistics
OR	Odds Ratio
OxIS	Oxford Internet Surveys
RAM	Random-Access Memory
RMSE	Root Mean Squared Error
RMSEA	Root Mean Squared Error of Approximation
RQ	Research Question
SAE	Standardised Absolute Error
SEM	Structural Equation Modelling
SM	Spatial Microsimulation
TAE	Total Absolute Error
TLI	Tucker-Lewis Index
TRS	Truncate-Replicate-Sample
UK	United Kingdom
UKHLS	United Kingdom Household Longitudinal Study
US	United States
VIF	Variance Inflation Factor
WHO	World Health Organization

1 Introduction

"Greater Choice and Control", "The Power of Information", "No Decision About Me, Without Me": the strategy tag lines employed by England's National Health Service (NHS) illustrate the transformational shift in the role of individuals in healthcare over the last decade (Department of Health, 2010, 2011, 2012a,b,c). Playing a central role in one's own health is considered to encourage healthy behaviours, promote self-care for chronic conditions, and improve judgement of when to seek professional help (Coulter, 2011). At the same time, transferring responsibility from the state to the individual may contribute to curbing rocketing healthcare costs in England and other nations (Santana et al., 2011).

The rise of the Internet has coincided with these developments, and may have contributed substantially to them (Powell and Boden, 2012; Anderson et al., 2003): while the medical knowledge base has grown significantly over the last century, the "sapiential authority" of doctors has declined due to the information available online (Gray, 2002, p. 6). Beyond factual information, individuals can network and share health insights online (Eysenbach, 2008), so that patients assume more active roles in their health (Broom, 2005) and are at the heart of digital health services (Tjora, 2014). With health systems under constant pressure to save money and improve quality, there are hopes that the Internet will help achieve both these aims through better informed patients and digital services which support the model of patient-centred care (Chaudhry et al., 2006; National Information Board (NIB), 2014).

Empirical research has shown that Internet use may indeed lead to health benefits through supporting informed decision-making and self-care (Rice, 2006), improving interactions with health professionals (McMullan, 2006; Wald et al., 2007), and providing social support (Alemi et al., 1996; Gustafson et al., 1999; Nabi et al.,

2013). Moreover, online resources may help to reduce unnecessary health service use (Gustafson et al., 1999), reduce emergency room visits (Dwyer and Liu, 2013), and increase help-seeking when appropriate (Berger et al., 2005; Suziedelyte, 2012). For society at large, the Internet also has the potential to level inequalities by reducing access barriers to care (Berger et al., 2005; Gilmour, 2007).

However, large-scale evidence on the health effects of Internet use is scant (Green and Rossall, 2013). Not everybody can and wants to assume responsibility for his or her own health (Henwood et al., 2003; Powell and Boden, 2012). Using the Internet may also have negative or no effects (Ziebland and Wyke, 2012; Eysenbach et al., 2004), partially depending on how it is used (Bessière et al., 2010), by whom (Eysenbach et al., 2004; Bell, 2014; Birru et al., 2004), and whether individuals can make sense of the quantity and quality of the information (Cline and Haynes, 2001; Sillence et al., 2007; Powell et al., 2003). "Cyberchondria" may potentially have negative effects on how individuals perceive their health, and lead to (unnecessary) use of health services (Loos, 2013). The real effects may be somewhere in between, as "both those espousing the empowering potential of the internet and those warning about its dangers overstate their cases" (Nettleton et al., 2005, p. 974).

The potential differences in individual outcomes also point to digital inclusion as an integral issue: to what extent are those without access to digital resources disadvantaged? Even though "inequality between 'haves' and 'have-nots' differentiated by dichotomous measures of access" may have been reduced with rising Internet penetration (DiMaggio and Hargittai, 2001, p. 2), inequalities remain in terms of the ability to make meaningful use of online resources (van Deursen and van Dijk, 2014) and to obtain benefits from doing so (van Deursen and Helsper, 2015). Since those with poorer health are often the most digitally disadvantaged (Dutton and Blank, 2013), it is likely that the benefits aspired to may not be realised, and that health inequalities are, in fact, exacerbated (Robinson et al., 2015).

Consequently, several UK/EU government initiatives aim to tackle digital inequalities in general (UK National Audit Office, 2013; European Commission, 2013) and for healthcare (Department of Health, 2013; European Commission, 2012). In 2013, the NHS commissioned a project with the name and aim of "Widening Digital Participation" (Tinder Foundation, 2013), an indication of the importance that the NHS places on closing digital divides to derive health benefits and increase efficiency. This mainly relates to the Internet as an information resource, though tackling digital exclusion may also "help address many wider equality, social, health and wellbeing issues such as isolation" (UK Government, 2014, no page).

1.1 Overview and research questions

In the light of the changed patient role, the potential benefits of the Internet for individual health and health systems at large, and the existing health and digital inequalities, this thesis addresses the following overarching research question (RQ): to what extent does Internet use influence individuals' health service use and their perceived health?

In this research, Internet use is defined as accessing the Internet on any device, given that the Internet is increasingly accessed via mobile devices (Wellman et al., 2003), but attention is also paid to the wider role of technology, particularly for Internet non-users (Rice and Katz, 2003). In addition to general use, this research also analyses specific online activities, particularly health information-seeking due to the outlined policy context. Following Kane and Radosevich (2011), the two health outcomes in this research are *perceived health* and *health service use*. Perceived health describes how individuals evaluate their health themselves (Blaxter, 2010). Health service use relates to traditional *offline* services, even though implications for online services will also be derived. Chapter 2 provides more information and detailed definitions of Internet use, perceived health and health service use.

More specifically, this research seeks to answer three interrelated questions in the context of Internet use and health:

RQ_1 To what extent does the Internet, specifically online health information, affect health service use and perceived health for Internet users and non-users?

RQ_2 What are the mechanisms by which Internet use affects health service use and perceived health?

RQ_3 How do individual and contextual factors influence the relationship of Internet use, health service use and perceived health?

While all RQs are linked to the main research question, each has a different purpose. RQ_1 relates to the general association between Internet use and health. It primarily addresses the area of digital inclusion by employing three user concepts derived from the digital divide literature: the *access divide* as "the gap between those who do and those who do not have access to computers and the Internet" (van Dijk, 2005, p. 1); the *support divide* in terms of the availability of proxies or other sources of help (Warschauer, 2004); and the *usage divide*, which relates to activities online (DiMaggio and Hargittai, 2001; van Deursen and van Dijk, 2014). RQ_2, focusing on mechanisms connecting the Internet and health service use/perceived health,

fills a theoretical gap by shedding light on *how* the two may be related, with core concepts of affordances (Gibson, 1986; Norman, 2002) and capitals (Bourdieu, 1986) connected to several health influence models from psychology and medical sociology (Andersen, 1968; Rosenstock, 1966; Mechanic, 1978; Suchman, 1965; Goldsmith et al., 1988; Uchino, 2004; Cohen, 1988). In contrast to the first research question, it moves beyond the Internet as an information resource and integrates mechanisms stemming from the Internet as a mediating and connecting technology. Finally, RQ_3 embeds this research in the wider individual and social context by analysing the extent to which sociodemographic factors and the social and geographic environment influence the relationship. This is theoretically relevant, with core concepts being Giddens' (1984) and Bourdieu's (1986) theories on structure and agency. Chapter 2 details how RQ_1–RQ_3 are derived from and situated in the literature.

To answer these questions, this thesis operates within the pragmatic paradigm and employs a mixed methods research design following a "quantitative – qualitative – quantitative" scheme. The quantitative part uses three datasets: the Oxford Internet Surveys (OxIS), the English census and Hospital Episode Statistics (HES) for 2011/13. These datasets are connected by spatial microsimulation (SM) based on output areas (OAs) in England. In SM, a simulated dataset is created from secondary data (existing surveys and routine health data) through probabilistic methods. As a spatial model, this provides a simulated dataset of everyone in a geographic area with their Internet use, health service use and perceived health, which then makes it possible to examine their associations through multiple regressions and structural equation models (SEM). To validate, enrich and contextualise the quantitative data, this research features 43 interviewees from the OAs employed in OxIS, with almost three quarters (31) former OxIS participants. This methodology allows for a detailed qualitative understanding of the mechanisms, while examining the larger picture of perceived health and health service use quantitatively.

The choice of these two health outcomes and their qualitative and quantitative exploration allows for a holistic perspective rooted in the World Health Organization's (WHO) definition of health as "a state of complete physical, mental and social wellbeing, and not merely the absence of disease and infirmity" (WHO, 1946, p. 2). Perceived health and health service use have close, mutually causal links (Johnson and Wolinsky, 1993; Kleinman, 1978; Andersen, 1995), as one partially predicts the other (see Chapter 2). Also, from a policy perspective, they need to be examined jointly: reduced health service use may be positive for reducing costs, but only if overall health status does not decline (Andersen, 1995).

This research focuses on England as a suitable environment for this investigation. As outlined, there are ongoing policy debates around the Internet in terms of healthcare and digital inclusion. The circumstance that the English population has largely universal access to healthcare via the NHS is an important feature for analysing health service use. From a practical perspective, data is available, with the NHS as a centralised health service and substantially wider computerised data capture than in the United States (US), Canada or Germany (Jha et al., 2008), the UK government being at the forefront of the open data movement (Davies, 2013), and OxIS as a random sample survey on Internet use for which contact details for follow-up interviews could be obtained for this research.

1.2 Background

With the focus of this research on England, it is important to understand the context of the NHS as the main way in which health services are provided to the population, and how the Internet is used for health purposes and in general.

1.2.1 Internet use in England

In 2013, about 79% of England's population used the Internet (OxIS, 2013). This number has steadily risen from 59% in 2003, showing that it is playing an increasing role in people's lives (Dutton and Blank, 2013). Broadband is available in 73% of households, and an additional 9% go online via 3G/4G on mobile devices. Challenges exist in rural areas in particular, with 3% of UK premises not being able to obtain even the lowest speed level of broadband (Ofcom, 2014).

Following Dutton and Blank (2013), more than two thirds of Internet users look up health information online, with a remarkable stability within the margins of the survey error (2007: 68%; 2013: 69%). Individuals across ages, educational backgrounds and income levels look up health information (Brodie et al., 2000), but the intensity of this varies: while the majority looks up health information monthly or less frequently (82% of health information-seekers), 14% do so weekly, and 4% daily (OxIS, 2013). Those with long-term conditions search for health information slightly more frequently (18% weekly; 4% daily), but tend to engage less in social activities (Dobransky and Hargittai, 2006). This pattern can also be observed for the elderly: while health information-seeking is relatively stable across age groups, social use declines with age. This particularly relates to social networking, as well as the producing and sharing of content (Dutton and Blank, 2013).

Then again, many of those with long-term conditions do not use the Internet at all, with disabilities and health problems being a "continuing source of digital exclusion" (Dutton and Blank, 2013, p. 22). While part of this difference disappears once socioeconomic status is controlled for, differences still persist (Dobransky and Hargittai, 2006). Following van Dijk and Hacker (2003), Internet access and use are unequally distributed along several dimensions: younger people use the Internet more than older people; higher education and socioeconomic status are correlated to Internet use; whereas the gender gap has nearly ceased to exist (Dutton and Blank, 2013). The exact boundaries between use and non-use may be blurry, as reflected in the shift towards referring to "digital inequalities" rather than to a single digital divide (DiMaggio and Hargittai, 2001; Warschauer, 2004). These inequalities comprise interrelated factors around skills, devices, attitudes and usage opportunities (van Dijk and Hacker, 2003; van Deursen and van Dijk, 2011a). Particularly in the health context, carers or family/friends often act as proxies for accessing health information online (Rice, 2006; Sadasivam et al., 2013), which again points to the importance of the social environment (Warschauer, 2004). Finally, these inequalities in Internet use also exhibit a geographic gradient, with digital exclusion highest in rural areas bordering Scotland and Wales, and lowest in and around London (Go ON UK, 2015; Blank et al., 2014).

1.2.2 Healthcare in England

England has a national health system financed through general taxation and run on the principles of equal access and being free at the point of care to ensure that access to health services does not depend on the ability to pay (NHS England, 2013). Individuals must register with a General Practitioner (GP), who is the main point of access for care and provides referrals to secondary care. In addition, services such as NHS 111 provide advice via the telephone, while NHS Choices (www.nhs.uk) makes it possible to check symptoms online and learn about various conditions.

Like many health systems globally, the NHS is under constant cost pressure (Crawford and Emmerson, 2012). At its inception in 1948, the assumption was that there was a fixed amount of illness which could be reduced by the NHS, which would lead to lower demand over time (Ham, 2009). This reflected the policy priority of increasing health service use in the middle of the last century (Andersen, 1995). However, the ageing population, new high-cost technologies, and a new generation of service users expecting higher standards of care has led to considerably higher demand (Ham, 2009): for example, there were 15.5 million hospital inpatient

admissions in 2013/14, 28% more than a decade earlier (Health and Social Care Information Centre (HSCIC), 2015; NHS Confederation, 2015). As a consequence, the NHS has struggled to maintain a stable budget while safeguarding quality of care and improving efficiency. The King's Fund (2014) calculates that even in the most optimistic scenarios, an additional two billion GBP in funding is needed annually, rising to eight billion per year by 2020.

Recent reforms have been designed to modernise the NHS by introducing payment by results, patient choice and performance transparency to drive efficiency and service orientation (Ham, 2009), with the Health and Social Care Act 2012 as the most significant change in the structure of health service provision. These developments have led to a more marketised NHS, with the idea of the empowered health consumer at its centre (Powell and Boden, 2012). For the increasing prevalence of Major and Chronic Diseases (MCDs), policymakers have emphasised the crucial role of self-management, reflected in the introduction of the Expert Patient Programme to train individuals in their long-term health conditions in 2001 (Department of Health, 1999, 2001). Beyond chronic conditions, the emphasis on information, choice and joint decision-making continues the trend of gradually shifting the burden of care from the state to the individual (Taylor and Bury, 2007).

Beyond the financial perspective, key concerns are improving quality of care and reducing health inequalities. Despite the universal access to healthcare, inequalities exist between groups with different socioeconomic status, education, and ethnicity (British Academy, 2014). Related partially to sociodemographic differences, there is also regional variation: general health declines from south to north (NHS England, 2010), and differs between smaller areas. Travelling east from Westminster, every two tube stations represent one year of life-expectancy lost; this is known as the "Jubilee line of health inequality" (London Health Observatory, 2010).

In response to the aims of reducing cost, improving care, and tackling inequalities, the NHS is increasingly counting on the benefits of information technology (NIB, 2014), in accordance with the "digital by default" vision for public services (UK Government, 2013). As previously discussed, the NHS envisions that patients "will have more choice and control, helped by easy access to the information they need" (Department of Health, 2010, p. 1). While technology has revolutionised medical care over the last decades, "the consumer experience of care services remains much as it was before the mobile phone and the internet became commonplace" (NIB, 2014, p. 8). Policymakers also highlight the convenience of the Internet, which may reduce the need for offline health services (Watt, 2015):

> *One thing missing from 111 is an online version of it. [...] I remember when one of my children had a rash. I think you would have been able to get a very good diagnosis very quickly if you had a computer screen that says, "which of these rashes does your child's most look like?" and that is something that is very difficult if you're talking to someone on the phone. [...] We have got plans to do it, [and] we definitely want it in the next couple of years, an online version of 111.*
>
> **– Jeremy Hunt**
> **Secretary of State for Health**

However, the use of digital services and the Internet as one of the main sources of health information (Pew Research Center, 2014; Dutton and Blank, 2013) rest upon the assumption that individuals can and want to engage online, and also presumes devices to access the Internet, skills, interest and usage opportunities (van Dijk and Hacker, 2003). Even though skills often dominate the focus, policymakers are aware of the need to enable access to digital offerings for everyone (NHS England, 2014b):

> *[We need to ensure that] people can access the support and information they need to improve their health and make informed choices. The fact that those most in need of NHS services are those least likely to have the skills they need to access online services is something that needs to be addressed urgently. The internet can be a powerful leveler in ensuring everyone can access the services that they need, but we need to invest in supporting those who don't have the skills.*
>
> **– Tim Kelsey**
> **Former National Director for Patients and Information**

To improve health, save costs and reduce inequalities, NHS England has made digital inclusion a priority, commissioning the previously mentioned "Widening Digital Participation" programme. The project trains 100,000 individuals, from hard-to-reach groups especially, to use the Internet with a focus on health-related activities (Tinder Foundation, 2014). Initial evaluations confirm that 93% of participants agreed they better understand how NHS Choices can assist them to manage their health, 76% felt more confident in using online health management tools, and 18% reported having reduced their use of offline NHS services as a result of accessing information online (Tinder Foundation, 2014).

However, it remains questionable whether health service use is indeed reduced, as Internet use may rather complement than replace offline services (Powell et al., 2011; Lupiáñez et al., 2013), and health professionals continue to play an integral role in individuals' health experiences (Cotten and Gupta, 2004; Henwood et al., 2003). This ties in with the wider question of how the Internet impacts health outcomes: is

it revolutionary due to fostering participation and democratising knowledge (Light, 2001; Forkner-Dunn, 2003), problematic due to potential misinformation and exclusion of certain population groups (Eysenbach et al., 2002; Robinson et al., 2015), or somewhere in between (Nettleton et al., 2005)? It is timely to better understand to what extent the Internet affects perceived health and health service use by proposing and testing theoretical pathways, as is done in this research.

1.3 Contributions of this research

By answering the RQs, and thereby shedding more light on the relationship between Internet use and health, the contribution that this thesis makes is three-fold: theoretical, practical and methodological.

1.3.1 Theoretical contribution

The first theoretical contribution is proposing and empirically testing mechanisms connecting Internet use, perceived health and health service use. Based on a lack of a conceptual framework in the literature, the contribution consists of bringing together "Internet research" from sociology and communications with models of health service use and perceived health from medical sociology and psychology. More specifically, this thesis connects affordances (Gibson, 1986; Norman, 2002) applied to the Internet (Wellman et al., 2003; Conole and Dyke, 2004), through which individuals may build up knowledge capital or social capital (Bourdieu, 1986; Savage, 2013), with models of perceived health (Uchino, 2004; Cohen, 1988) and health service use (Andersen, 1968; Rosenstock, 1966; Mechanic, 1978; Suchman, 1965; Goldsmith et al., 1988). This research also integrates elements of social cognitive theory (Bandura, 1986, 2004) to explain how the Internet may shape health behaviour, which in turn affects both outcomes of perceived health and health service use.

By analysing the access divide, support divide and usage divide (DiMaggio and Hargittai, 2001; Warschauer, 2004; van Dijk and Hacker, 2003; van Deursen and van Dijk, 2014) with a specific focus on health, this thesis also contributes to the digital inclusion literature. With its focus on health outcomes, it sheds light on a potential "outcomes divide" (also referred to as a beginning third-level divide by van Deursen and Helsper, 2015), as Internet-based outcomes have generated high scholarly interest in recent years (Eysenbach, 2003; Rice, 2006; Helsper, 2012; van Deursen and Helsper, 2015). By drawing on the "uses-and-gratifications" approach, this research also incorporates antecedent factors around technology adoption and

use (LaRose and Eastin, 2004), and responds to the criticism that "hardly any substantial empirical or theoretical effort has been devoted to connecting gratifications and effect" (Katz et al., 1973, p. 518).

Finally, this thesis adds an understanding of the influence of individual and contextual factors rooted in the longstanding debate on the interplay of agency and structure, or micro and macro effects (Turner, 2005). Drawing on Giddens' (1984) structuration theory as the theoretical and methodological base, this research also integrates Bourdieu's (1986) work with concepts such as *fields* and *habitus*. Rather than treating these influences purely as factors that need to be controlled for, this thesis develops insights about the role of individual sociodemographic factors known to shape both Internet use and healthcare, as outlined in section 1.2, the role of offline social circles (Haythornthwaite and Wellman, 2002), and area-based influences on perceived health and health service use (Kawachi and Berkman, 2003).

1.3.2 Practical contribution

From a practical point of view, the findings of this research may contribute to the debate around quality and efficiency in the health system, with the two outcomes – perceived health and health service use – touching on both of these aspects. Understanding the health-related effects of Internet use and developing appropriate, targeted measures helps policymakers and practitioners to take advantage of the upsides of technology-enabled healthcare while reducing potentially adverse effects. While this research focuses on offline health services, it also has implications for the online provision of health services in a more digitally connected health system.

For the policy debate around digital inclusion and health, this research sheds light on the different (health-related) facets. While current policy documents by the NHS and the Department of Health (UK Government, 2013, 2014) already evidence an appreciation of the topic, it is important to go beyond devices and the frequently mentioned skills challenge (see Tim Kelsey, p. 18), and to understand factors around interest and the perceived usefulness of the Internet (van Dijk and Hacker, 2003). With this, this research may also inform the ongoing NHS project "Widening Digital Participation" and further programmes aimed at training (and potentially motivating) people to use the Internet in order to benefit their health.

At the same time, the SM technique allows to break down the results to smaller areas, thereby allowing for an understanding of differences by location, such as urban versus rural, or translating findings into more specific area-based policy actions (Tanton and Edwards, 2013). There has been continued interest in understanding

the implications of Internet use and health at the level of local authorities, such as Solihull (Anderson and Speed, 2009, 2010) or Cambridgeshire, as evidenced in a practitioner workshop on digital inclusion in February 2015. Moreover, this research also provides tangible levers for action by segmenting effects according to different kinds of individuals, and analysing how sociodemographic factors and their social and geographic environment play in the equation.

Finally, doctors do not universally welcome patients coming in with a stack of printouts obtained online (Lewis, 2006); particularly when the Internet started to become more commonplace, some have been found to act in a defensive way when faced with patients bringing in health information obtained online (Murray et al., 2003). To this end, this research also delivers insights for practitioners with respect to the diverse ways in which health-related Internet use influences patients' decisions and needs. This is particularly important to support the positive effects of Internet use while reducing its potentially adverse implications.

1.3.3 Methodological contribution

From a methodological point of view, SM is an "innovative way of combining diverse datasets in order to understand health and wellbeing" and enables "high levels of exploitation of our existing data investments" as set out in the Economic and Social Research Council's strategic priorities (ESRC, 2009, p. 7; ESRC, 2015a, p. 16). At the same time, this research links to the priorities of the Medical Research Council to better understand the complex relationship of lifestyle, inequalities and other measures from outside the healthcare setting (MRC, 2014). Through SM, aspects that would have traditionally been considered less relevant in health research can be incorporated, including Internet use and the action of various social conditions causing disease (Link and Phelan, 1995). While Whitacre and Brooks (2014) have already used area-based data to understand links between broadband adoption and health, SM allows for the incorporation of individual sociodemographic covariates.

SM has been used in other health-related contexts, such as analysing social media use (Anderson and Speed, 2009, 2010), evaluating access to GP services (Morrissey et al., 2008), understanding patterns of depression (Morrissey et al., 2010, 2015), and estimating smoking prevalence (Tomintz et al., 2008; Smith et al., 2011). The purpose of these studies was to estimate the geographic distribution of certain variables and use them for subsequent analyses, whereas this research primarily employs SM as a technique for linking datasets, using geography as the connecting element rather than an end in itself. This research thereby demonstrates a novel

way of applying SM to examine the relationship between Internet use and health through existing secondary data, which otherwise could not be analysed jointly. In addition, this study enriches the body of health-related SM literature with its focus on Internet use as a more recent area of interest in health research. Finally, while smaller SM models – for a city or a local government district – provide very detailed insights (so that "size" is not necessarily delivering better results), the large-scale focus on the whole of England, combined with insights broken down to the local level, increases applicability across contexts and questions of policy relevance.

Another contribution is showing the value of the mixed methods research design. As opposed to the majority of SM studies, this research adds a qualitative perspective, and thereby enriches and confirms the insights from the simulated dataset based on qualitative interviews with some of the original OxIS participants. A simulated dataset based on probabilistic methods will not be as accurate as obtaining measures from everyone in the population: actual microdata may still be seen as the gold standard (Rees et al., 2002), as are surveys which can be collected more frequently than the census and yet provide good estimates if correctly sampled (Fowler, 2014). However, SM is a methodologically relevant avenue in cases where no existing survey covers the questions under consideration, and, particularly when combined with qualitative data, provides a way to explore connections between concepts that are frequently not contained within the same dataset.

Finally, this research also shows how existing data allows for the derivation of insights in a privacy-preserving way (Hermes and Poulsen, 2012). This is crucial in an era in which more data is produced than can potentially be analysed, particularly as health research is invariably linked to privacy and data protection. Reusing secondary data is also a cost-effective way of creating new insights. Against a backdrop of finite researcher resources, this leaves more time for other activities (Teddlie and Yu, 2007), for example for the qualitative data collection in interviews.

1.4 Thesis structure

Building on the overview of this research, its background and contributions, Chapter 2 reviews the theoretical and empirical literature, derives the three RQs, presents the analytical framework, and systematically details each of its parts in its four sections. Section 2.1 reviews existing models of health service use and perceived health (for example, Andersen, 1968; Rosenstock, 1966; Cohen, 1988), which provide the connecting points for various Internet-based effects. As section 2.2 explains,

these effects may relate to different user concepts (van Dijk, 2005; DiMaggio and Hargittai, 2001; Warschauer, 2004; van Deursen and van Dijk, 2014), and individuals' motivations and enablers (LaRose and Eastin, 2004), which shape not only Internet access and use, but also its outcomes. Section 2.3 focuses on *how* the Internet, its affordances (Gibson, 1986; Norman, 2002) and its capitals (Bourdieu, 1986; Savage, 2013; Grossman, 1972) link to the identified influence factors on perceived health and health service use. Finally, section 2.4 discusses the effects of individual factors, such as age, gender, education, socioeconomic status and long-term conditions, and contextual factors, such as the social and geographic environment.

Chapter 3 – the methodology chapter – explains the choice of the pragmatic research paradigm, and the consequential rationale for a mixed methods design to examine the relationship between Internet use and health (section 3.1). For the quantitative part, section 3.2 explains the selection of all datasets, the SM process, and the internal and the external validation of the resulting dataset. For the qualitative part, section 3.3 details the two-stage sampling process of areas and individuals, and explains the interview guide. The quantitative and qualitative approaches are reunited in section 3.4, which outlines the analytical strategy used to derive insights from both parts of the mixed methods design jointly and separately. The chapter ends with ethical considerations in section 3.5.

Chapter 4 presents the results for RQ_1, disentangling the effects of online health information on perceived health and health service use for Internet users and non-users. Starting with the main descriptive results for both health outcomes in section 4.1, section 4.2 analyses their levels for the three user concepts based on the digital divide literature. Based on different motivations and enablers for health information-seeking identified in the qualitative research, as described in section 4.3, section 4.4 presents a typology of orientations towards health information online (learners, pragmatists, sceptics, worriers, delegators and adigitals). Revisiting the quantitative data in section 4.5 shows that the typology is useful for understanding Internet-based effects on health service use and perceived health.

The second results chapter on mechanisms of use (Chapter 5) focuses on Internet users to understand the underlying mechanisms of Internet use, as RQ_2 set out to do. This chapter draws primarily on the qualitative data to enable a detailed understanding of the mechanisms at work, both for health service use (section 5.1) and perceived health (section 5.2). Where possible, given the restricted nature of quantitative data, section 5.3 revisits the quantitative data to model some of the observed mechanisms. While this chapter focuses primarily on Internet users, some

qualitative non-user accounts are included to enrich the wider picture. At the same time, Chapter 5 extends the typology developed in Chapter 4 by differentiating between these mechanisms by types where applicable.

Chapter 6, the final results chapter, analyses influences on Internet use, the two health outcomes in this research, and their relationship (RQ_3). Section 6.1 starts with individual factors such as age, education, socioeconomic status and long-term health conditions, providing insights from the qualitative data followed up by quantitative analysis. In the same manner, section 6.2 presents the results for the social environment. Finally, section 6.3 addresses the spatial dimension, drawing on the quantitative data first and using the qualitative data to explain the results. It analyses compositional and contextual differences between rural and urban areas, as well as regional variations. The chapter ends with a case study of the Isle of Wight as a specific local authority district.

The discussion in Chapter 7 reunites the findings from the three results chapters. It starts out by linking the observed mechanisms to the analytical framework in section 7.1, which also serves as an overview of theoretically relevant topics detailed in the following sections. For the area of digital inclusion, section 7.2 describes the shift towards an outcomes divide, outlines the role of the social environment as a catalyst and filter, discusses parallel mechanisms in the offline world, and suggests the "lifestyle paradox". Building on these insights, section 7.3 discusses how the Internet may influence structures that enable and restrict agency (Giddens, 1984) in the context of health. Section 7.4 shifts the focus to implications for practice and policymaking, organised along major policy debates in England: enabling outcomes while leaving nobody behind; shifting responsibility from the state to the individual; and enabling data-driven insights for healthcare. To conclude, section 7.5 discusses the chosen mixed methods research design and the suitability of both strands to answer the RQs under consideration, and suggests that qualitatively informed constructs be added to the secondary data used in this research in future waves.

Chapter 8 concludes by synthesising the findings for RQ_1–RQ_3 in section 8.1. Section 8.2 then summarises the contributions of this research to the body of theoretical work, its practical value and recommendations for policy, and the benefits and drawbacks of the methodology employed. The thesis concludes in section 8.3 with directions for further research in the area of Internet use and health.

2 Literature Review

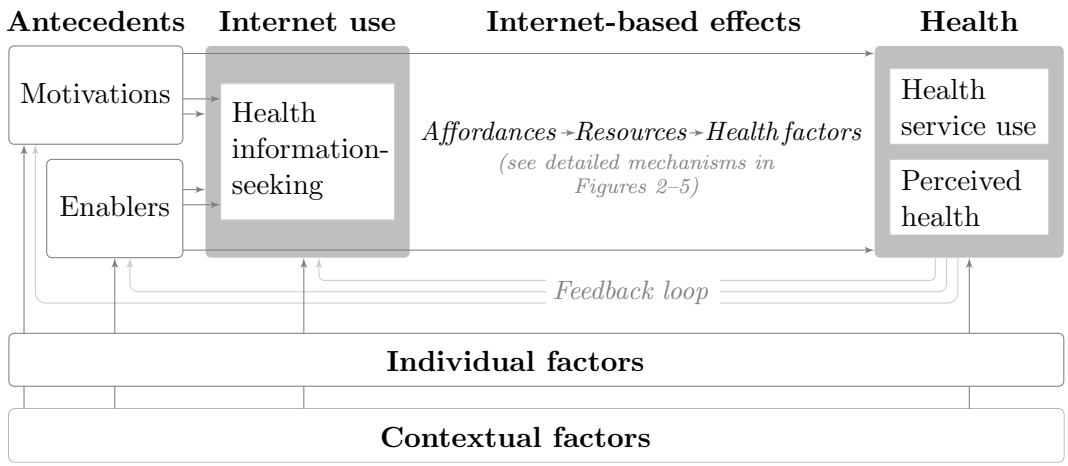
A survey of word use in written English between the 1930s and today showed that those that gained most dramatically in frequency are "around", "health", "information", "it's", "didn't", "says", "social", "family", "children" and "need" (Baker, 2011, p. 74). While these do not necessarily stand in conjunction, their increased usage suggests that health and health information, as well as individuals' needs and their social environment, count among the topics of growing interest during the last century. The academic literature relevant to this research has contributed to their increased occurrence: researchers from various disciplines have empirically examined aspects of Internet use and health throughout the last two decades, while the relevant theoretical foundations from (medical) sociology, psychology and communications reach back to the middle of the last century.

Consequently, this chapter has three main purposes interwoven throughout all its sections: first, it reviews the more recent empirical literature about Internet use and health outcomes. Appendix I (p. 287) presents an overview of studies identified through a structured literature review (Petticrew and Roberts, 2006). It includes all original research about the relationship of Internet use (including connected software or apps) on either perceived health or health service use published in peer-reviewed journals or books. It is important to relate these to the Internet's changing role in everyday life (Haythornthwaite and Wellman, 2002) and geographic differences, for example between the US and the UK, in terms of access to and financing of health services. These caveats will also be highlighted, where applicable, when reviewing the empirical literature, particularly in sections 2.3 and 2.4.

The second purpose of this chapter is underpinning the relationship between Internet use, perceived health and health service use with stronger theoretical foundations than are currently provided by the empirical literature. To do so, this research draws on health-related models on health service use (Andersen, 1968; Kleinman, 1978; Mechanic, 1978; Suchman, 1965; Goldsmith et al., 1988), perceived health (Cohen, 1988; Uchino, 2004) as well as social cognitive theory and its applications to the health context (Bandura, 1986, 2004). These models are linked to theory on the effects of Internet use, integrating the uses-and-gratifications approach (Katz et al., 1973), media affordances (Gibson, 1986; Norman, 2002) and their application to the Internet (Wellman et al., 2003; Conole and Dyke, 2004), and capitals built up through them (Bourdieu, 1986; Grossman, 1972). Internet use is conceptualised following the access, support and usage divide derived from the digital divide literature (van Dijk, 2005; DiMaggio and Hargittai, 2001; Warschauer, 2004; van Deursen and van Dijk, 2014). For the wider frame of the micro level of the individual and the macro level of their social context, this research draws on Bourdieu's (1986) and Giddens' (1984) work on the duality of agency and structure.

Integrating these theoretical concepts, this thesis builds and empirically tests an analytical framework on the relationship between Internet use, health service use and perceived health. The third purpose of this chapter is therefore to introduce and gradually develop the analytical framework proposed in this research shown in Figure 1. It posits that general and health-related Internet use have affordances, by which individuals may build up resources, which in turn affect previously identified influence factors on perceived health and health service use in various ways. These mechanisms are complex and will be detailed step-by-step in Figures 2, 3, 4 and 5. Internet use in turn is influenced by motivations and enablers, which may also influence outcomes. However, as captured by the feedback loop (Andersen, 1995), health itself may influence Internet use and its antecedent factors: a health problem may act as a motivation for health-related Internet use (Dwyer and Liu, 2013; Xiao et al., 2014), while it may also prevent individuals from using the Internet (Dobransky and Hargittai, 2006; Houston and Allison, 2002). Individual and contextual factors (age, gender, education, socioeconomic status, long-term conditions, as well as the social and geographic environment) influence and potentially mediate these relationships. The analytical framework provides the foundation for building and analysing the SM model and developing the interview guide in Chapter 3, directs the analysis in Chapters 4, 5 and 6, will be revisited in the discussion in Chapter 7, and is presented in an updated form in Chapter 8.

FIGURE 1: Overview of analytical framework



The analytical framework rests upon the argument that "explaining Internet behavior entails understanding that 'the Internet' is not a separate entity, but instead a complement to ongoing activity", which we can only understand by "considering a more integrated view of people's lives" (Haythornthwaite and Wellman, 2002, p. 35). While the individual is central to and the unit of analysis in the empirical part, the social and local context is integral as well (Rosenstock, 1966; Andersen, 1995). Existing frameworks (Eysenbach, 2003; Dutta-Bergman, 2006) have the same intention, but only relate to health outcomes as abstract constructs, disregard established models of perceived health and health service use from the pre-Internet era, and either neglect mediating mechanisms (Dutta-Bergman, 2006) or individual and contextual influence factors (Eysenbach, 2003). Further insights from both frameworks will be discussed in the literature review where applicable.

The remainder of this chapter systematically examines each part of the analytical framework, reviewing the theoretical and empirical literature jointly and explaining how both tie into the analytical framework. Section 2.1 presents the two health outcomes of perceived health and health service use, showing that both are influenced by various factors to which the Internet may contribute. Section 2.2 then focuses on the effects of Internet use, also introducing the three user concepts from the digital divide literature and antecedent factors around motivations and enablers of Internet use, and leads to RQ_1. Section 2.3 looks at *how* the Internet may influence health service use and perceived health, reviewing the Internet's affordances and capitals built up through them to derive RQ_2. Finally, section 2.4 discusses the influence of individual and contextual factors, and thereby provides the basis for RQ_3.

2.1 Health

This research employs perceived health and health service use as the two primary health outcomes (Kane and Radosevich, 2011). In line with the centrality of the individual to the inquiry, these outcomes reflect the patient-centric perspective, as opposed to clinical evaluations from a medical professional, for example. They do not relate to any specific health condition, which compromises their specificity (Kane and Radosevich, 2011), but allows for a holistic view of the patient's health experience and increases their relevancy for deriving implications for policy and practice (Bowling, 2009; Andersen, 1995).

Perceived health describes how individuals subjectively rate their current health status, and is a "surprisingly useful measure [to distinguish] the medically defined healthy or unhealthy" (Blaxter, 2010, p. 24). While perceived health is associated with actual health and long-term conditions (Goldstein et al., 1984; Hunt et al., 1980; Fillenbaum, 1979), both may also deviate: for example, older individuals may include age-related problems in their evaluations (Schneider et al., 2004; Kivinen et al., 1998; Leinonen et al., 2002), and those with severe health problems who score low on objective health assessments may still perceive their health as very good, as captured by the disability paradox (Albrecht and Devlieger, 1999). Perceived health therefore adds an important perspective beyond biological measures, as it incorporates social, psychological and physical elements of health (Bowling, 2009; Wilson and Cleary, 1995). Notably, it has shown a high predictive value for mortality in longitudinal research (Idler and Benyamini, 1997; Kaplan and Camacho, 1983; McGee et al., 1999), to the extent that it predicted mortality better than the medically determined health status (Mossey and Shapiro, 1982), and has yielded particularly good validity when predicting help-seeking and health service use (Branch et al., 1981; Johnson and Wolinsky, 1993; Miilunpalo et al., 1997; Dening et al., 1998; Bowling, 2005).

Health service use is another patient-centric, albeit more objective, health-related outcome that refers to if, when and how often individuals receive professional medical help in primary or secondary care for mental and physical conditions (Kane and Radosevich, 2011). As opposed to perceived health, health service use should not necessarily be high or low, but rather appropriate (Rosenstock, 1966). This may differ on a case-by-case basis (Hannay, 1980), with changing policy priorities between periods of time: while the aim in the middle of the last century was to increase and enable health service use, current considerations about cost and efficiency rather seek ways of reducing it (Andersen, 1995; Ham, 2009, see section 1.2). As outlined

in the introduction, health service use refers to traditional offline services in this research. While Internet use likely affects online consultations, and these may in turn affect offline consultations (Bergmo et al., 2005), online consultations are only coming into existence at the moment (NIB, 2014). Nevertheless, implications of this research for online health services will be included in the discussion due to its potential importance for the future of health service provision.

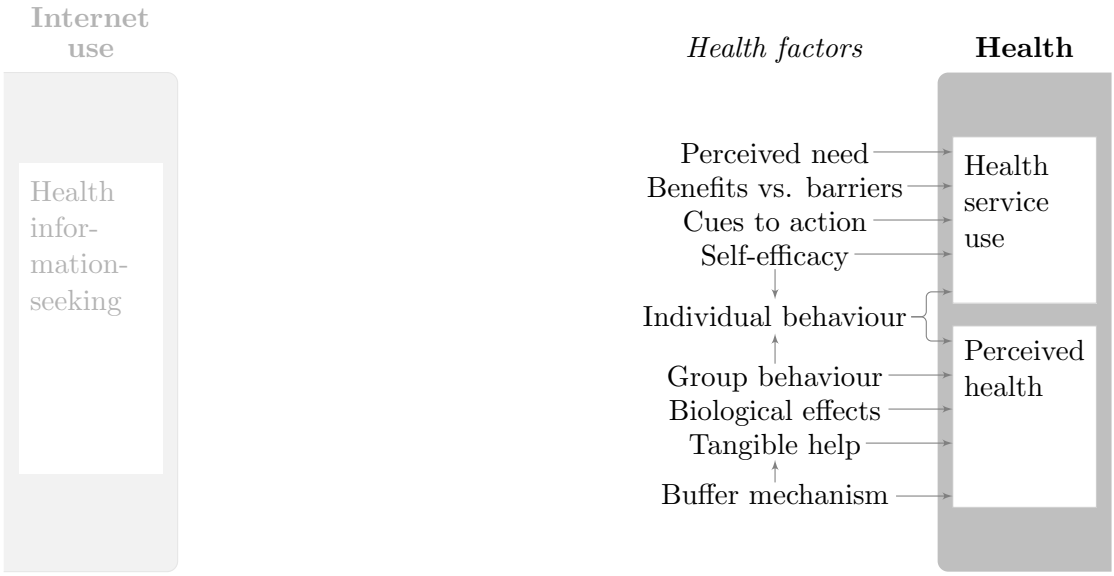
Both perceived health and health service use are quantifiable outcomes that entail the practical corollary of being accessible via a single question (Bowling, 2005) or from existing records (Kane and Radosevich, 2011). This is in line with the aspiration of this research to understand the Internet's effects on the frequency of health service use and individuals' health rating. Yet this research does not treat these outcomes only in a quantifiable sense; equally important are the less measurable dimensions around decision-making processes about health service use, and the nuances of health perceptions that cannot be captured by ticking boxes or choosing a number on a scale. This makes mixed methods approaches a particularly fruitful approach for exploring these outcomes, as will be detailed in Chapter 3.

Perceived health and health service use cannot be examined in isolation from each other due to their mutual dependency and joint practical relevance (Andersen, 1995; Kleinman, 1978, see section 1.1). Even with different conceptualisations and measurements, how individuals rate their own health has "consistently been shown to be the best predictor of the use of health services" (Johnson and Wolinsky, 1993, p. 105). It is still worth examining them separately: utilisation of services has often been (mis-)used in the literature as a proxy for health, but differences in access (which may be less pronounced in the context of England) and other cultural and economic factors may influence utilisation (Kane and Radosevich, 2011). Additionally, those who see medical specialists frequently do not automatically suffer from poor health, while those with poor health do not always access the services they should. Hannay (1979, 1980) use the metaphors "iceberg" and "trivia" to suggest how health services may be over- or underutilised: while many medically relevant conditions remain "below the surface" and are never brought to medical attention, medical help is sometimes also unnecessarily sought for trivial ailments.

The question of necessity can be evaluated only by including a professional's perspective, and will not be covered by the patient-centric view assumed in this research. The focus of this research, and specifically the remainder of section 2.1, is rather to identify what influences health service use and perceived health, with the aim of establishing a discrete set of factors to which the Internet may contribute.

Figure 2 introduces a high-level view of these factors and how they fit into the larger analytical framework, which will be completed throughout Figures 3, 4 and 5.

Figure 2: Analytical framework (1/4) – influences on health



The remainder of this section explains these influences on health service use (subsection 2.1.1) and perceived health (subsection 2.1.2), and how they were derived from the various theoretical models. As with Bandura's (2004, p. 145) observation for health behaviours, many models "include overlapping determinants but under different names, [and] facets of a higher order construct are often split into seemingly different determinants [...] following the timeless dictum that the more the better". With the primary intention of finding the key factors influencing both health outcomes, the terminology has been harmonised as far as possible.

2.1.1 Models of health service use

Widely employed models of health service use date back to before the rise of the Internet, and include the sociobehavioural model of health service use (Andersen, 1968, 1995), Rosenstock's (1966) psychological model of health service use, Mechanic's (1978) general theory of help-seeking from a psychology background (incorporating insights from Mechanic, 1972, 1975), Goldsmith et al.'s (1988) help-seeking decision-making model, and Young's (1981) choice-making model (as cited in Rebhan, 2009). As Rebhan (2009) and Wyke et al. (2013) suggest, health service use may rather need to be conceptualised as an iterative process, understanding interrelated sequences

of disease events and taking into account previous experiences, rather than evaluating singular decisions. Disease process models relevant for understanding health service use include Suchman's (1965) illness model, Kleinman's (1978) outline of the different (non-)professional spheres in which decisions about health service use are made, the illness action model (Dingwall, 2001), the common-sense model of the self-regulation of health and illness (Leventhal et al., 2003), and Pescosolido's (1992) model rooted in a social organisation strategy approach.

Across these models, there are three common influences on health service use, which led to the first three factors in the analytical framework in Figure 2. The first cluster is around perceived need, which is closely linked to the perception of symptoms and their salience, frequency and persistence, their perceived seriousness, and the associated disruption of normal activities (Andersen, 1968; Mechanic, 1972, 1978; Rosenstock, 1966; Leventhal et al., 2003). The second cluster is around expected benefits from the use of health services and their convenience (Rosenstock, 1966; Young, 1981; Suchman, 1965) relative to their barriers in the sense of effort, but also stigma and denial of illness (Rosenstock, 1966, 1974; Mechanic, 1978). Third, there may be specific cues to action: these may come internally as symptoms, but also externally from others, through reading information, particularly through media, or from hearing about others who are affected (Rosenstock, 1966; Rosenstock et al., 1988; Goldsmith et al., 1988; Pescosolido, 1992; Dingwall, 2001). Appendix II (p. 291) contains an overview of all models.

In addition, several individual factors affect health service use, often referred to as predisposing characteristics or health beliefs (Rosenstock, 1966). Health beliefs are attitudes and values about the expected benefits of health services, barriers to them and individuals' needs, and may also be influenced by how people "experience and emotionally respond to symptoms of illness, pain, and worry about their health condition" (Andersen and Davidson, 2007, p. 8). Sociodemographic variables also play a role: age affects the likelihood that individuals need health services, while education and socioeconomic status may influence the ability to cope with problems and command resources to solve them (Andersen and Davidson, 2007). These factors will be detailed in sections 2.2 and 2.4.

Nearly all models also highlight contextual factors, particularly the social environment, as "most decisions regarding when to seek aid [...], whom to consult, and whether to comply, along with most lay evaluations of the efficacy of treatment" (Kleinman, 1978, p. 86) are made outside of formal medical encounters. Social contacts may help individuals to notice and evaluate symptoms (Mechanic, 1975;

Goldsmith et al., 1988; Dingwall, 2001), while social pressure and group influences may affect help-seeking (Goldsmith et al., 1988; Pescosolido et al., 1998; Suchman, 1965), for example through providing cues to action (Rosenstock, 1966), shaping health beliefs (Pescosolido, 1992), and contributing knowledge (Dingwall, 2001).

While all of these models address health service use, several have developed to include health behaviour more generally. For example, Rosenstock's (1966) model of health service use was extended to the outcome "likelihood of action", which includes engagement in health-promoting behaviour (Janz and Becker, 1984; Rosenstock, 1974, p. 334). A decade later, it was enhanced through elements from social cognitive theory (Rosenstock et al., 1988; Bandura, 2004). Similarly, Andersen and Davidson (2007) extended Andersen's (1968, 1995) health service use model with behavioural elements such as "personal health practices" and "process of medical care". A more detailed review of these health behaviour models is beyond the scope of this research. Nevertheless, health behaviours will be given due consideration, and are included in the analytical framework (see Figure 2), as they are important predictors of both health service use and perceived health (Andersen, 1995).

Bandura's (1986) social cognitive theory and his application of it to health (Bandura, 2004) provides a useful addition to the introduced models for three reasons. First, it posits "knowledge" as a crucial determinant of health behaviours, which is also reflected in the process-oriented models about health service use (Wyke et al., 2013; Dingwall, 2001; Leventhal et al., 2003; Pescosolido, 1992) and serves as a link to the effects of Internet use in section 2.3. Second, a key element for translating knowledge into action is self-efficacy (Bandura, 2004), which relates to the belief that certain goals can be achieved. Following Bandura (2004), self-efficacy helps to determine aspirations for health behaviours, whether individuals stick to their commitments, and what agency they assume for their health (see section 2.4). In line with Eysenbach (2003), self-efficacy is therefore added to the analytical framework in Figure 2 as a separate influence on health service use. Third, in line with many models on health service use (Andersen, 1995; Dingwall, 2001; Pescosolido, 1991; Mechanic, 1975; Young, 1981; Goldsmith et al., 1988; Rosenstock, 1966), social cognitive theory highlights the importance of the social environment for acquiring knowledge and behaviours. For media influences, Bandura (2004, p. 150) compares the direct pathway, in which "media promote changes by informing, modeling, motivating, and guiding personal changes", with the socially-mediated pathway, in which media links people, who then provide each other with guidance and support for behavioural changes. This duality is also reflected in the inclusion of individual

and contextual factors for Internet use (Hale et al., 2010) and health (Riva et al., 2009) in section 2.4. What Bandura's (2004) theory adds is that the pathways by which the two health outcomes are influenced may also be mediated by these factors.

Finally, it should be kept in mind that the presented models largely focus on illness, as opposed to the WHO's wider definition of health employed in this research (see section 1.1). In addition, while some models primarily viewed health service use as going up or down based on these influences, it is important "to view it as dynamic and part of a broader system mediated by patient's life world experience of chronic illness and trajectories of utilisation", and acknowledge the interrelated nature of these influence factors (Gately et al., 2007, p. 944). Health service use is influenced by "illness careers" (Wyke et al., 2013, p. 80) and individuals' general history (Dutta-Bergman, 2006). As defined in section 2.1, health service use will therefore be seen as a broader concept, including the individuals' experiences (to which the Internet may contribute), previous encounters with the health system, which may shape motivations for health information-seeking, and related health perceptions, which will be detailed in the next subsection.

2.1.2 Models of perceived health

Health and an individual's perception thereof are also shaped by multiple interrelated factors. The most salient influence may be around individuals' impairments, with chronic conditions, both physical and mental, often leading to lower evaluations of health (Johnson and Wolinsky, 1993; Hansen et al., 2001). In addition, individual factors such as age and education (Johnson and Wolinsky, 1993; Kawachi et al., 1999), and predisposing personality characteristics (Goodwin and Engstrom, 2002) may shape health perception. Similarly, individuals' social context plays a role (Fillenbaum, 1979), with Durkheim's (1952) work from 1897 on social integration and suicide rates representing early work in this area despite its methodological challenges (Berkman et al., 2000).

The literature has not always differentiated between actual and perceived health. Biomedical research has primarily investigated actual health, including cardiovascular disease, dementia and mortality (Holt-Lunstad et al., 2010; Cacioppo et al., 2002; Knox and Uvnäs-Moberg, 1998; Fratiglioni et al., 2000), while research from sociology, psychology and epidemiology has also included health perception (Wan, 1976; Garrity et al., 1978; Schneider et al., 2004). Both are relevant, as perceived health may influence morbidity and mortality (Idler and Benyamini, 1997). This link may be explained based on Bandura's (2004) self-efficacy, which is central to shaping

health behaviours and improving health (Duncan and McAuley, 1993; Grembowski et al., 1993), and Antonovsky's (1979) similar concept of sense of coherence, which suggests that an optimistic outlook and feelings of control positively affect health (Schneider et al., 2004; Suominen et al., 2001). This research subsumes both actual and perceived health, but primarily addresses *social* influences, which serve as connecting points for the Internet's effects in sections 2.2 to 2.4.

In contrast to health service use, the literature on perceived health is relatively fragmented. However, Cohen (1988), Berkman et al. (2000) and Uchino (2004) systematised the main mechanisms stemming from the social environment, particularly social support (see Appendix III, p. 294). Social support may differ in terms of structural features (the kind of networks) and functional differences (what individuals use it for, Berkman et al., 2000). In its widest definition, it includes informational, emotional, psychosocial and tangible support (Cohen, 1988; Reblin and Uchino, 2008). Following Cohen (1988), perceived health may be influenced by four major factors: information to ease or evaluate ailments and create behavioural change; group behaviours that shape individual behaviour; psycho-/biological factors; and the provision of tangible resources. Each of these represent factors in the analytical framework (Figure 2) and will be detailed below.

First, information from a certain source or from within the individuals' social network links back closely to the relationships discussed for health service use, evidencing the interrelated nature between both outcomes. As discussed for Bandura's (2004) social cognitive theory, information may lead individuals to engage in health-promoting behaviours, reduce stressors and health risks, with a healthier lifestyle having positive influences on health (Cohen, 1988). This rests upon the assumption that this information is beneficial, as it may also lead to the discouragement of service use or inadequate alternative treatments, or to the promotion of unhealthy behaviours (Cohen, 1988). Consequently, information may shape individuals' perceived health, but this influence may be either positive or negative.

This two-sided nature also exists for the second influence factor: others' influence on health behaviour. Certain habits may be viewed with disapproval (smoking, drugs) or encouraged (exercise, healthy food) by the social environment, thereby positively affecting health (Treiber et al., 1991; Reblin and Uchino, 2008; Lindström et al., 2000). This also includes norms and values, as "the social influence which extends from the network's values and norms constitutes an important and underappreciated pathway through which networks impact health" (Berkman et al., 2000, p. 849). Observing other people may lead an individual to adopt their behaviours

(Bandura, 1986), and these in turn may influence psychological and biological processes that link to health (Cohen, 1988). The beneficial effects depend on the kinds of health behaviours promoted through the social environment, although potentially harmful habits were found to contribute little to individuals' health perception when controlling for sociodemographic covariates (Ericsson, 1997).

Third, the social environment may play a psycho-/biological role in influencing health, as "no matter how well informed patients are about their medical condition, most need the reassurance of talking to others who have been in similar situations, who can demystify the health delivery system, and who can provide emotional and social support" (Alemi et al., 1996, p. 33). Social support can create or amplify feelings of self-esteem, self-identity, and recognition of self-worth, all of which may be associated with better health (Cohen, 1988). Researchers have hypothesised underlying cardiovascular, endocrine, immune system-related, and inflammatory processes (Uchino et al., 1996; Uchino, 2006), and argued that social isolation as "a chronically stressful condition to which the organism responds by aging faster" (Berkman, 1988, p. 51) also has negative health consequences (Eysenbach, 2003).

Fourth and finally, tangible support provided by the social network may result in higher health perception. Beyond the importance of having someone available when ill, which positively affects health perception (Rose, 2000), tangible support includes help with various daily mundane tasks, as well as access to resources and material goods, such as jobs and economic opportunities, and access to healthcare (Berkman et al., 2000). Of course, there are close links between all the different influence factors on perceived health, given that "tangible and informational support may also serve an emotional support function, as when they signal caring and are not viewed as resulting from obligation" (Schaefer et al., 1981, p. 386).

The causal pathways between these four factors and perceived health are complex and disputed (Cohen and McKay, 1984; Cohen and Wills, 1985; Alloway and Bebbington, 1987): direct or main effects models posit that the factors mentioned above influence health directly, while buffer or stress-centred models see social support as a moderator (Cobb, 1976; Cassel, 1976), suggesting the "intuitively acceptable idea that being able to turn to people for support mitigates the effects of misfortunes" (Alloway and Bebbington, 1987, p. 91). This dispute may largely stem from the inconsistent use of social environmental effects in previous studies (Cohen and McKay, 1984), reflecting the distinction between structural and functional features of social support previously mentioned. Cohen and Wills (1985) concluded that buffer effects exist when social support captures the availability of help to respond to

the needs created through stressful life events, whereas direct effects may be observed when social support rather relates to individuals' integration in social networks. In this research, both approaches are taken into account, as both potentially play a role for understanding the mechanisms by which Internet use affects perceived health.

Finally, the mere *perception* of available support may come with health benefits (Schaefer et al., 1981; Wethington and Kessler, 1986). Even *giving* support has also been associated with better health outcomes – either because individuals benefitted from feelings of social usefulness, or because those giving more support also received more support (Reblin and Uchino, 2008; Wang et al., 2015). However, support for the situation at hand may not always be available (Cohen and McKay, 1984; Berkman et al., 2000), may be inappropriate, and may create feelings of reciprocity, dependency, conflicts and "relational strain" (La Gaipa, 1990; Due et al., 1999, p. 663) which potentially negatively affect health.

While this section has discussed influence factors on health service use and perceived health, both may also influence individuals' daily activities as well as their Internet use. For example, a health problem or a physician consultation may motivate a search for health information (Dwyer and Liu, 2013; Xiao et al., 2014; Caiata-Zufferey et al., 2010), while health problems may also impede necessary enablers, either due to physical constraints, or due to concomitant circumstances such as lower financial resources (Dobransky and Hargittai, 2006; Houston and Allison, 2002). These "feedback loops" (Andersen, 1995, p. 7) will be discussed further throughout the following section.

2.2 Internet use

This section now shifts the focus from health to Internet use, presenting both the user concepts employed for analysing the Internet's effects, and explaining factors antecedent to Internet use. However, what "Internet use" is has become increasingly blurred. Various devices and locations to access the Internet, and apps and messaging services on mobile devices, make the boundaries of Internet use less clearly defined from the user's point of view. In addition, the individual's social context is invariably linked to access (Warschauer, 2004): some non-users may access the Internet via proxy, while some users may need others' help. As included in section 1.1, Internet use is defined here as personally using any device for any online activity in any location (Dutton and Blank, 2013). Particularly for non-users, this research also pays attention to the wider social and technological context, including mobile phone use and proxy access (Warschauer, 2004; Rice and Katz, 2003).

2.2.1 Three user concepts based on the digital divides

The distinction between use and non-use has also been captured by the digital divide literature. With increasing Internet penetration, the terminology has shifted towards referring to "digital inequalities" (DiMaggio and Hargittai, 2001), or distinguishing between nuanced divides: the *access divide* in terms of whether individuals access the Internet; the *support divide* with respect to proxy use and the availability of help; and the *usage divide* about online activities (van Dijk and Hacker, 2003; Warschauer, 2004; DiMaggio and Hargittai, 2001; van Deursen and van Dijk, 2014). The user concepts in this research are derived from these digital divides, as they capture theoretically and practically important aspects beyond the user/non-user dichotomy. Table 1 shows their definitions and proportions in England based on OxIS (2013).

Table 1: User concepts following three digital divides

	Access divide	**Support divide**	**Usage divide**
User (79%)	Next-generation user (54%) *user with access on mobile and/or several devices*	Independent user (36%) *user who worked things out without help*	Health user (55%) *user who seeks health information online*
	First-generation user (25%) *all other users with access, on laptops/desktops, for example*	Supported user (26%) *user who regularly receives others' help*	Non-health user (24%) *user who does not seek health information online*
		Unsupported user (17%) *user who did not work out things and received no help*	
Non-user (21%)	Ex-user (3%) *non-user who used the Internet in the past*	Proxy access (4%) *non-user who asked somebody for Internet use*	
	Never-user (18%) *non-user who has never used the Internet*	Proxy availability (11%) *non-user who knows someone, but has not asked*	
		Fully excluded (6%) *non-user with nobody to use the Internet for them*	

Source: The access divide follows Dutton and Blank (2011); see Appendix IV (p. 295) for variable definitions. Percentages refer to the population of England.

The **access divide** in terms of having a device connected to the Internet was the primary way to distinguish use and non-use in the past (van Dijk, 2005). However, it would be erroneous to believe that "information inequality [...] is solved at the moment that everyone has the ability to obtain a personal computer and a

connection to the Internet" (van Dijk and Hacker, 2003, p. 316), as interest, skills and usage opportunities also matter (van Dijk, 1999, see subsection 2.2.4). Then again, the access divide still remains relevant, as also the experience of connectivity influences Internet-based outcomes (Allen, 2010). More recently, the kind of device may also fall within the scope of the access divide. Dutton and Blank (2011) coined the term "next-generation users" for those going online on mobile or multiple devices. While first-generation users also use the Internet for health information, next-generation users are more likely to see the Internet as a convenient first port of call, create content and interact with others online (Dutton and Blank, 2013).

However, the "pressing question will be *not* 'who can find a network connection [...] from which to log on?', but instead, 'what are people doing, and what are they *able* to do, when they go on-line?'" (DiMaggio and Hargittai, 2001, p. 3–4, emphasis in original). This links to the **support divide**, as Internet use is invariably linked to the social context, both in terms of getting help for users, and providing proxy access for non-users (DiMaggio and Hargittai, 2001; Warschauer, 2004). Across contexts, social networks were among the most crucial determinants of access, activities, and outcomes of Internet use (Helsper, 2008; Eynon and Malmberg, 2012). In particular, online health information-seeking "is a team sport" (Sadasivam et al., 2013, p. 193), and the help of "warm experts" (Bakardjieva, 2005, p. 99) has not only proven crucial for technical questions, but also for finding and interpreting information (Wyatt et al., 2005; Rice, 2006).

DiMaggio and Hargittai's (2001, p. 4) reference to "what are people doing [...] when they go on-line" also relates to the **usage divide** (van Dijk and Hacker, 2003; van Deursen and van Dijk, 2014). The usage gap rests upon the assumption that certain activities are more beneficial than others to build up resources and achieve outcomes (van Deursen and van Dijk, 2014, see section 2.3). In this research, usage primarily relates to health information-seeking, which forms part of everyday life information-seeking activities that "people employ to orient themselves in daily life or to solve problems not directly connected with the performance of occupational tasks" (Savolainen, 1995, p. 267). This focus is useful due to the importance of knowledge for perceived health and health service use (see section 2.1), information-seeking being a frequent activity across users (Dutton and Blank, 2011), and online health information in principle being accessible to both Internet users and non-users.

This conceptualisation of Internet users reflects two interrelated debates in the theoretical literature. The first is the relationship between digital and social inclusion, as the most socially deprived are also least likely to use online resources

(Helsper, 2008). For users, social inclusion relates to the ability to engage meaningfully online (Warschauer, 2004), while for non-users, the social context may act as a proxy to access online health information (Rice, 2000). However, strong social networks may also decrease the willingness to engage online due to the convenience associated with accessing information via others (Reisdorf et al., 2012), showing the multifaceted links between digital and social inclusion.

The second underlying theoretical question involves the shift in the digital divide. Even though access, support and usage are still relevant concepts today, their relative importance may gradually change: individuals' breadth of online activities differs just like the access divide around the turn of the millennium, with usage following the same kind of socioeconomic patterns characterising access found then (van Deursen and van Dijk, 2014; van Deursen and Helsper, 2015). Following van Deursen and van Dijk (2014), this may primarily reflect the wider context of use in terms of motivations and enablers for use. Similarly, Rogers and Mead (2004) found that how individuals use health information online is largely shaped by their perceptions of technology, previous health service use, and their belief in the usefulness of health information, all of which are detailed in the next subsection.

2.2.2 The importance of antecedent factors

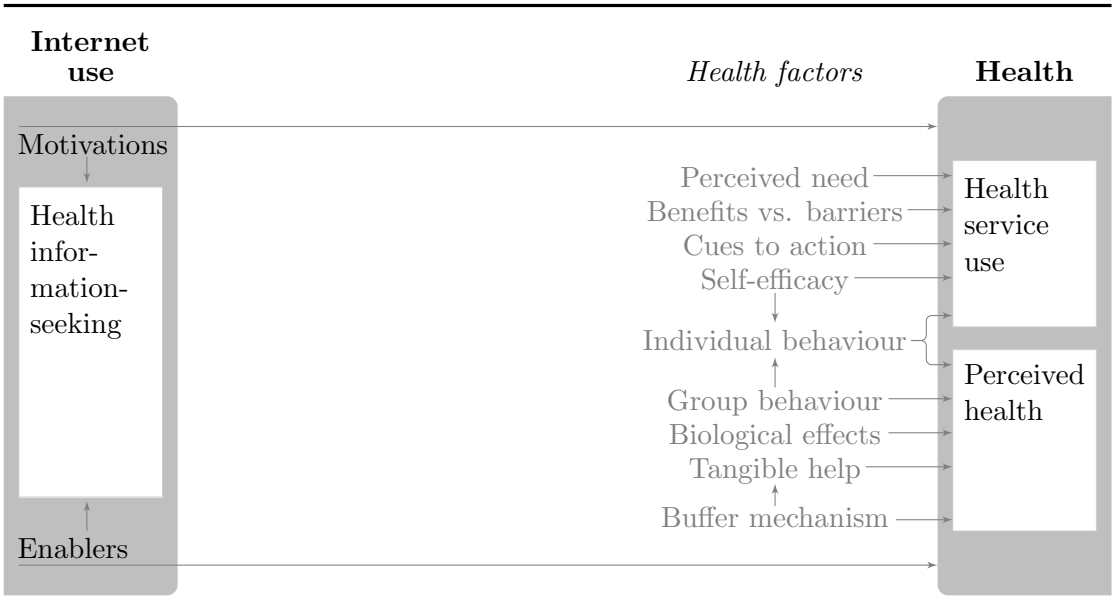
Supporting the importance of antecedent factors around motivations and enablers for the uses and effects of the Internet (van Deursen and van Dijk, 2014; Rogers and Mead, 2004), Powell et al. (2011, no page) concluded that the Internet creates "evolutionary rather than revolutionary change", as online health resources become part of normal health practices, help-seeking, and everyday life information-seeking (Nettleton et al., 2005; Kivits, 2009; Savolainen, 1995). Anderson and Tracey (2008, p. 160) summarised this by saying that "people are not doing anything particularly new, they are doing old things in new ways and finding that some of those ways suit their lifestyles better." This ties in with the wider argument that individuals or households use technologies to satisfy existing needs, with technological innovation merely creating new ways of doing so (Gershuny, 1983), as also reflected in theories around the social shaping of technology (MacKenzie and Wajcman, 1999). Particularly for health, online resources may have reinforced the development of some patients questioning medical authority and assuming more active roles (Gray, 2002). Without resolving the wider debate on the bidirectional influences of technology and society (see Fuchs, 2008), the point is that antecedent factors of Internet use must be taken into account when analysing use and non-use (Reisdorf and Grošelj, 2015).

Incorporating these antecedent factors in the analytical framework is important as they do not only affect use, but also the outcomes of it (Helsper, 2012; Straubhaar et al., 2012). People's "disposition to technologies [...] interact in a reciprocal and complex relationship" with how they build up resources online (Straubhaar et al., 2012, p. 8), also because the ability to do so is mediated by access, skills and attitudes (Helsper, 2012). This relates to technical skills and interest (van Deursen and van Dijk, 2011b), but also to health literacy as the "capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions" (National Library of Medicine, 2014, p. VI).

In addition, motivations and enablers may also directly influence health outcomes. For example, low health literacy is associated with poorer health outcomes (Parker, 2000) due to differences in self-management, lifestyle, use of preventative service and response to health education (DeWalt et al., 2004). Consequently, and related to the previous discussion of technology in fact reinforcing existing motivations (Powell et al., 2011; Nettleton et al., 2005; Kivits, 2009; Anderson and Tracey, 2008), motivations for health information-seeking may be congruent with wider motivations to exhibit agency and a desire to be involved in health, which are important influences on perceived health and health service use (Rosenstock, 1966).

Consequently, motivations and enablers are included as influences on Internet use, health information-seeking, and health outcomes (Dutta-Bergman, 2006), but also "in between" as a technology mediator, as shown in Figure 3.

Figure 3: Analytical framework (2/4) – enablers and motivations



The remainder of this section further develops the antecedent factors presented. Subsection 2.2.3 frames motivations for general and health-related Internet use through a uses-and-gratifications approach (Katz et al., 1973; Stafford et al., 2004; LaRose and Eastin, 2004), before subsection 2.2.4 conceptualises its enablers based on the digital divide and digital inclusion literature (van Dijk and Hacker, 2003; Warschauer, 2004; DiMaggio and Hargittai, 2001; van Deursen and van Dijk, 2014).

2.2.3 Motivations for health-related Internet use

Extending van Deursen and van Dijk (2014), this research conceptualises motivations for general and health-related Internet use based on individuals' expected gratifications. Uses-and-gratifications approaches posit that individuals choose media based on their expected returns, which in turn are influenced by individuals' psychological characteristics, their social environment, their expectations, and the assumed consequences (Rubin, 2009). Originating in the middle of the last century, uses-and-gratifications approaches marked the shift from guiding research questions of "what do the media do to the people?" to "what do people do with the media?"(Katz, 1959, p. 2) and "sought to replace the image of the audience member as a passive victim [...] with one of a person who could actively bend programmes, articles, films, and songs to his own purposes" (Blumler, 1979, p. 10).

The more active choice by the media user conceptually matches the idea of patients becoming more active in their healthcare (see section 1.2), particularly in terms of the Internet with its unprecedented interactivity (LaRose and Eastin, 2004). Media choice can be explained by various approaches, and what medium people use also depends on their motivations (Dutta-Bergman, 2006). However, in the context of this research, the uses-and-gratifications approach primarily serves as a frame for understanding antecedent factors and their effects on health service use and perceived health. By linking gratifications to outcomes, this research also contributes to an area that was often neglected in the early days of uses-and-gratifications research (Katz et al., 1973), although the contemporary research focus has shifted towards understanding media effects (Rubin, 2009).

Media use gratifications can be derived from three sources: the content that the media allows access to, media exposure itself, and the social context which the medium is embedded in (Katz et al., 1973). These three sources of gratifications have been confirmed conceptually and empirically for the Internet (Stafford et al., 2004; LaRose and Eastin, 2004). However, some manifestations of these gratifications are slightly different for the Internet: opportunities for active learning and finding

information on the content side; activities around searching and surfing related to the medium per se; and interacting with others online among the social connectivity gratifications. However, particularly for the last aspect, one may argue that there is "an entirely new social gratification that is unique to Internet use" (Stafford et al., 2004, p. 259), depending on what medium the Internet is compared to.

Katz et al. (1973) argued that content gratification has been the basis of research from the very beginnings, with less attention being paid to gratifications from media use per se and social gratifications (Katz et al., 1973). While health information is still the centre of attention in the literature on health and the Internet (Powell et al., 2005), and uses-and-gratifications approaches have been applied to health information-seeking in recent publications (Rains, 2007; Tustin, 2010), this is not quite true today anymore. Research has also looked at how anonymity, convenience of access online, and support from people are the key benefits that individuals ascribe to the Internet (Powell et al., 2011; Tang and Lee, 2006).

The motivations for health-related Internet use found in previous research focus on looking for reassurance for a condition or treatment; wanting to get a second opinion to challenge previous information; trying to understand or supplement previous information or doctor's advice; attempting to overcome barriers to accessing information via other channels (Powell et al., 2011; Ybarra and Suman, 2006; Chapple et al., 2012); and, particularly for long-term conditions, tackling isolation and finding social connections (Ziebland et al., 2004). Reasons for general Internet use in principle relate to the same dimensions in a less specific and fine-grained way. For example, employing a uses-and-gratifications approach, Papacharissi and Rubin (2000) identified interpersonal utility, passing time, seeking information, convenience and entertainment as motivations for general Internet use. Moreover, health information-seeking shows similar usage patterns as other information-seeking activities, as it falls within the same principal component (Blank and Grošelj, 2014).

In the Internet context, uses-and-gratifications approaches have been theoretically extended in various directions. LaRose and Eastin (2004) suggested increasing the explanatory power through incorporating elements from Bandura's (1986) social cognitive theory, given that self-efficacy also shapes what gratifications individuals expect online. As a common element with influences on perceived health and health service use (see section 2.1), this further supports the idea that motivations may be linked to Internet use and health outcomes, and potentially to their relationship. Other relevant constructs for analysing expected gratifications and Internet use are experience with the Internet and online behaviours (LaRose and Eastin, 2004).

2.2.4 Enablers for health-related Internet use

In order to understand Internet-based outcomes, motivations and expected gratifications must be viewed in conjunction with whether individuals actually possess the ability to realise them (Dutta-Bergman, 2006). Enablers for accessing online resources may be clustered around four themes: interest in going online; Internet-enabled devices and a connection to the Internet; skills for doing so; and relevant usage opportunities (van Dijk and Hacker, 2003). DiMaggio and Hargittai (2001) delineate comparable enablers, but in addition relate to autonomy of use and social support. These are already reflected in the three user conceptualisations employed in this research, and also point to the importance of both individual and contextual elements in the social environment (Warschauer, 2004). In addition, taking up LaRose and Eastin's (2004) point about Internet experience again, "across all studies, the largest and most significant differences in access and use are related to years of experience" (Haythornthwaite and Wellman, 2002, p. 17).

These prerequisites to using the Internet are highly interrelated. For example, those without devices may not have a device because they are not interested, and may not build up the required skills – with different causal pathways between all of these aspects. In addition, individuals require various skills: using the technology (also referred to as "computer literacy" or "computer-mediated communication literacy"), but also finding and evaluating content online (Warschauer, 2004; van Deursen and van Dijk, 2011a). In the context of health information, this also relates to defining search terms, selecting the best sources, and evaluating health information (van Deursen and van Dijk, 2011b), encompassed in the concept of health literacy (mentioned above). Finding information can be challenging, especially for those with low Internet literacy or health literacy (Birru et al., 2004). Achieving positive health benefits through online sources also depends on the ability to deal with information (Broom, 2005), apart from the general link between low health literacy and poor outcomes related to socioeconomic indicators (Parker, 2000). As a consequence of the theoretical and empirical evidence, enablers will also be taken into account to understand how Internet use affects health outcomes.

2.2.5 RQ_1: Health information for (non-)users

Based on this review of Internet use and its antecedent factors around motivations and enablers, the first RQ addresses the larger picture of the effects of Internet use on perceived health and health service use. It seeks to understand different

kinds of users and non-users, utilising user concepts following the access, support and usage divide (see Table 1). At the same time, as individuals may fall into the categories of these user concepts for different reasons, the first RQ goes beyond this view and incorporates individuals' motivations and enablers, which may both shape Internet use and its effects (Helsper, 2012; van Deursen and van Dijk, 2014; van Deursen and Helsper, 2015). Contributing to the digital inclusion literature and to understanding the Internet-based outcomes, the first results chapter (Chapter 4) therefore addresses the following question:

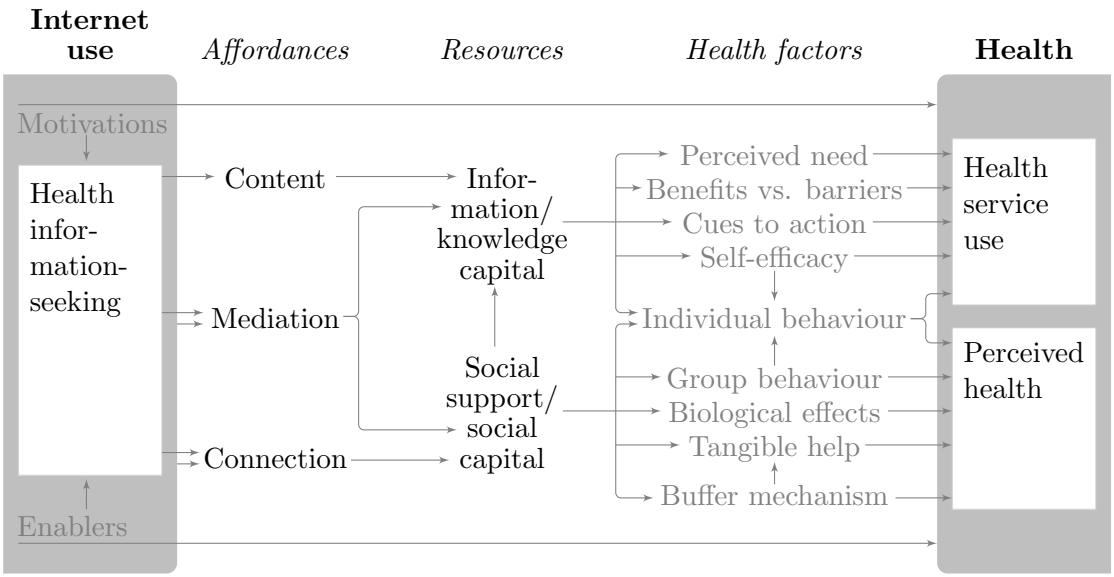
RQ_1 To what extent does the Internet, specifically online health information, affect health service use and perceived health for Internet users and non-users?

At the same time, the user concepts and the role of antecedent factors call for a more detailed understanding of the processes by which Internet use, perceived health and health service are related. In several instances, this section has already highlighted the importance of activities online, and the idea that individuals build up resources through them – depending on their motivations in terms of *wanting* to engage in their health, and their enablers in terms of *being able* to. Extending and complementing these arguments, these potential mechanisms between Internet use, perceived health and health service use will be detailed in the next section.

2.3 Effects of Internet use on health

This section now turns to the mechanisms by which the effects of Internet use may be connected to the factors influencing health service use and health perception outlined in section 2.1. Its main argument is that Internet use has affordances (Gibson, 1986; Norman, 2002) through which individuals may build up capital resources (Bourdieu, 1986), and which then shape perceived health and health service use. Among the different kinds of capitals, the analytical framework in particular focuses on knowledge capital and social capital, through which information and social support may be exchanged. This choice rests upon the suggestion in the previous literature that both knowledge capital and social capital may be built up online (Savage, 2013; Wellman et al., 2001), while they also link to the influences on perceived health and health service use described in section 2.1. The updated framework is shown in Figure 4, with each part being detailed in the remainder of this section: affordances in subsection 2.3.1, and the resources derived from it in subsections 2.3.2 (information/knowledge capital) and 2.3.3 (social support/social capital).

Figure 4: Analytical framework (3/4) – affordances and resources



2.3.1 Affordances: Content, mediation and connection

Internet use may have a range of affordances, a term that has been defined inconsistently in the literature. Gibson (1986) describes affordances of the environment as "what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill" (p. 127, emphasis in original). He argues that the central point of affordances is not whether they are real, but whether they are apparent for an individual. The "environment" is rather fuzzily defined in Gibson's (1986) theory of affordances: he refers to the environment in general, but also to different places, while "the richest and most elaborate affordances of the environment are provided by other animals and, for us, other people" (Gibson, 1986, p. 135). Following Gibson, affordances may be both beneficial and injurious, exist without extensive learning by the individual, and may be interpreted as being binary: either they do or do not exist for a given combination of an individual and the environment (McGrenere and Ho, 2000). In contrast, Norman's (2002) definition focuses on individuals and objects rather than the general environment, adopting primarily a design perspective and therefore seeing affordances as properties of objects. In a similar vein, and in contrast to Gibson (1986), Norman distinguishes between real and perceived affordances, and attributes an important role to the past knowledge and experience that individuals apply to understand an object's affordances (Norman, 2002).

Both Gibson's and Norman's concept of affordances has been used throughout the technology affordances literature, sometimes unintentionally mixing both (McGrenere and Ho, 2000; Gaver, 1991). However, in terms of understanding the Internet's effects, both contain relevant elements. First, Norman's allowance that affordances can be influenced by knowledge and experience may be better suited to capture the idea of enablers, as the Internet's affordances may not be intuitively apparent to everyone. Second, the complexity of user concepts described in subsection 2.2.1 contrasts with Gibson's notions of binary affordances. Unifying both Norman's and Gibson's position, the view of affordances as dependent on the individual's perception links in with the uses-and-gratifications approach, suggesting that affordances depend on the individual's motivations. Finally, Gibson's concept of relating both to other individuals and places better captures the idea of a social technology that enables contact with other people and places. For the purpose of this research, affordances are defined as perceived ways in which the Internet can be used by an individual depending on knowledge and experience.

Several scholars have discussed affordances of the Internet. Wellman et al. (2003) primarily focused on the social affordances of Internet use, and enumerated five: broader bandwidth, a permanent connection, personalisation, wireless portability and globalised connectivity. Similarly, Conole and Dyke (2004) created a long list of learning-related affordances: accessibility, speed of change, diversity, communication and collaboration, reflection, multimodal and non-linear, risk, fragility and uncertainty, immediacy, monopolisation and surveillance. Common to the literature is that it explicitly or implicitly shows a mix of Gibson's and Norman's affordance concepts, and extends the meaning of affordances quite significantly: for example, "broader bandwidth" describes a technical feature of the Internet that provides the basis for affordances such as different forms of content or interaction. Moreover, the existing literature either focuses on certain aspects of affordances, for example social affordances (Wellman et al., 2003), or provides long unstructured lists of general affordances (Conole and Dyke, 2004).

To systematise the affordances of the Internet, this research extends the uses-and-gratifications approach: with expected gratifications relating to the content accessed through the medium, the medium per se and the social context or connections enabled through the medium (Katz et al., 1973; Stafford et al., 2004, see subsection 2.2.3), the Internet's affordances may also follow these categories. Table 2 shows the affordances of the Internet from this perspective, with some of Wellman et al.'s (2003) and Conole and Dyke's (2004) being summarised or worded differently.

Table 2: Affordances of Internet use

Content	Mediation	Connection
– Instant information – Information anywhere, anytime (day/night) – Active, personalised information searching – Diversity of content	– Constant communication (always on) – Enhanced experience (with pictures and videos) – Cue reduction (for example, written communication only) – Asynchronicity – Mass communication and crowdsourcing – Geographical reach	– Daily coordination with friends/family close by – Keeping in touch with friends/family farther away – Making new connections (for example, related to common interest)

Source: Wellman et al. (2003) and Conole and Dyke (2004)

Out of these three categories, mediation is particularly new, as the Internet creates mediation affordances that may not have existed before. The other two affordances are primarily redefined in the context of the Internet. For example, medical information has been available in magazines and newspaper articles or pamphlets at the doctor's surgery, but actively seeking this information was time-consuming, while the information may also have lacked the variety and, where applicable, went out of date at some point (Anderson et al., 2003; Broom, 2005). Similarly, while the connection affordance in principle was available via the telephone, finding new ties in particular may be enabled by the Internet.

While these affordances relate to general use, they also apply in a health-related context, as the Internet "provides convenient and anonymous access [to health information] at any time, from any location, to a wide range of expert sources; and through virtual communities it can provide peer support and social interaction" (Powell et al., 2011, no page). Similarly, Jeremy Hunt's quote (p. 18) highlighted several affordances of the Internet: information is readily available at one's fingertips, and pictures may help identify a rash, for example, by visual comparison.

2.3.2 Resources: Information and knowledge capital

These affordances of the Internet may help to build up a range of resources (Helsper, 2012; Hampton and Wellman, 2003), which represents the second step in the proposed analytical framework (see Figure 2). Resources are conceptualised here in the Bourdieusian (1986) sense as long-lasting capacities "to produce profits and to reproduce itself in identical or expanded form" (p. 46). Bourdieu (1986) highlighted that

there are different forms of capitals – such as social and cultural capital – thereby extending the primary form of economic capital that dominated the literature up to this point. Even health may be conceptualised as a capital good (Grossman, 1972), in the sense that individuals have a certain stock at the beginning of their lives that depreciates with age, and which may be sustained by health service use and certain health behaviours.

The different forms of capital are not independent of one another. Following Bourdieu (1986), different forms of capital may be converted into others, so that for example, health capital may be turned into economic capital (healthy people face fewer work-related restrictions than those with a long-term health condition). At the same time, economic capital may also shape health service use, as the time spent on going to the doctor's may be more valuable for obtaining additional resources (Grossman, 1972). This foreshadows the link to the relevance of individuals' characteristics, which will be discussed in section 2.4.

A particular resource that may be built up online is around information, which may also be conceptualised as knowledge capital due to its long-lasting value for decision-making (Savage, 2013). Knowledge capital may be considered a form of cultural capital following Bourdieu (1986), but for the health-related context in this research, and due to the role of knowledge in health service use and perceived health models (see section 2.1), it is useful to refer to knowledge capital specifically. Of course, one may particularly speak of knowledge *capital* for those with long-term conditions, often referred to as "expert patients" or "lay experts" (Prior, 2003; Ziebland, 2004), which is reflected in the Expert Patient Programme (Department of Health, 2001). For ad hoc requests, it may be more appropriate to speak of information that individuals obtain when needed (and may forget afterwards).

Effects of information and knowledge capital vary. On the positive side, the "expert patient" is associated with improved health outcomes through empowerment for better self-care and joint decision-making in the doctor-patient relationship (Powell et al., 2003), through better adherence to physician recommendations (Bowling, 2009; Samal et al., 2011), or through seeking help when needed (Suziedelyte, 2012; Berger et al., 2005). Even if information does not influence decisions, it may improve subjective wellbeing and perceived health (Broom, 2005). However, as discussed for motivations and enablers (see section 2.2) the ability to achieve health outcomes partially depends on the ability to deal with information (Broom, 2005), and again on education – with low health literacy being consistently related to lower health outcomes and increased health service use (Berkman et al., 2011, see section 2.4).

Then again, information may also be wrong or even harmful (Rice, 2006). While diversity is positive, navigating the enormous quantity of information online may pose difficulties (Powell et al., 2003). In addition, health information online may create negative emotions (Lauckner and Hsieh, 2013), increased anxiety and rumination (Bessière et al., 2010) and, with information available 24/7, "concerns and anxieties that may be difficult to deal with in the middle of the night" (Ziebland and Wyke, 2012, p. 224). Moreover, individuals may misinterpret information, follow inappropriate treatment advice, or incorrectly question conventional treatments (Cline and Haynes, 2001; Broom, 2005), thereby reducing health service use with potentially negative consequences. However, "cyberchondria" may increase health service use (Loos, 2013), reinforced by the tendency of search engines to push up concerning information to the top of the search results (White and Horvitz, 2009).

To a certain extent, the real effects of health-related Internet use may be somewhere in the middle (Nettleton et al., 2005), which is reflected in the empirical literature that specifically analysed the relationship between Internet use and health (see Appendix I, p. 287). Most studies found moderate impacts on either of the two outcomes, and none reported any substantial harm (Crocco et al., 2002; Fox, 2011). While there may be a slight publication bias since all studies confirmed most of their initial hypotheses, this supports the idea of a middle ground between revolutionary effects and serious harm. For example, for health service use, early research by Gustafson et al. (1999) and Alemi et al. (1996) discovered more efficient and appropriate health service use, while Rice (2006) found no effects. In contrast, Berger et al. (2005), Suziedelyte (2012), Lee (2008), Ybarra and Suman (2006) and Dwyer and Liu (2013) found a positive correlation between looking up information online and health service use, particularly for stigmatised conditions (Berger et al., 2005), although urgent visits and emergency services declined (Dwyer and Liu, 2013).

For perceived health, Rice (2006) confirmed positive effects of health information-seeking on perceived health, particularly after health service use (Ybarra and Suman, 2006), whereas Bessière et al. (2010) and Alemi et al. (1996) found no effects on this outcome. Interestingly, for broadband availability in small areas, Whitacre and Brooks (2014) confirmed higher levels of perceived health, but with the interesting observation that "broadband adoption rates are negatively associated with the percentage of people who describe themselves as in very good health, while they are positively associated with the percentage of those describing themselves as in good health" (p. 774), which points to the complex relationship between Internet use and health. Similarly, for depression as a specific health outcome, Bessière et al.

(2010) found that informational use led to higher levels of depression, whereas social use improved depression. Although neither changed perceived general health, this shows the importance of incorporating both informational and social pathways, as will be done in the next subsection.

2.3.3 Resources: Social support and social capital

The Internet may also help to build up social capital, through which information and social support can be exchanged. There are multiple definitions of social capital (Adler and Kwon, 2000), but in this research, it is defined as the "aggregate of the actual or potential resources which are linked to possession of a durable network" (Bourdieu, 1986, p. 51). This primarily captures the idea that social capital facilitates the acquisition of resources through connections with others (Wellman and Frank, 2001), in contrast to, for example, Putnam's (2000) participatory capital, which concerns involvement in politics and voluntary organisations (Wellman et al., 2001). In that sense, social capital is the infrastructure through which social support – as one of the predecessors of the factors influencing perceived health (see section 2.1) – may be exchanged. More specifically, "network members provide emotional aid, material aid, information, companionship, and a sense of belonging," so that social support is the main way to "obtain resources to deal with daily life, seize opportunities, and reduce uncertainties" (Wellman and Frank, 2001, p. 1).

Social capital may be built up through the Internet in various forms, reflecting the connection affordances of the Internet (see subsection 2.3.1): through finding new weak ties online, for example in online support groups, but also through enriching contact with existing strong ties, such as family and friends. For the former, individuals may build up networks of similar-minded individuals, or individuals with similar concerns (Wright and Bell, 2003). For strong ties, Wellman et al. (2001) concluded that Internet use supplements previous social capital, as it extends face-to-face contact and telephone conversations, although this also depends on what people do online (as reflected in the usage divide).

Social capital may also be particularly enabled based on the mediative affordances of the Internet. Due to cue reduction online (for example, exchange of text only, without eye contact or oral communication), the Internet enables a unique social distance and "hyperpersonal" communication to create feelings of social support (Walther, 1996, p. 3). This rests upon the early Internet logic of acting as a medium to create new weak ties, providing wider access to people interested in particular topics. Weak ties facilitate anonymous "low-risk discussions about high-risk topics"

(Adelman et al., 1987, p. 135), so that the Internet creates social augmentation (Bessière et al., 2008; Kraut et al., 2002). Particularly in the context of health, the Internet created connections to strangers who could be asked embarrassing questions (Hardey, 1999), and getting their help in online support groups was associated with experiencing higher levels of social support (Fogel et al., 2002).

Nevertheless, some researchers have argued that the Internet damages social contacts, thereby decreasing social support. The central argument is that it is "strong social ties [...] that generally buffer people from life's stresses and that lead to better social and psychological outcomes" (Kraut et al., 1998, p. 1019), and by using the Internet, people displace social activity or substitute poorer quality weak ties for better quality strong ties (Kraut et al., 1998). This research was carried out during the early days of the Internet in everyday life, and did not incorporate other factors, such as interest in the Internet and offline social circles, which might influence whether relationships are displaced or augmented (McKenna and Bargh, 1999). Empirical evidence shows that existing offline social support is important for the effects of Internet use (Kraut et al., 2002), also in the specific context of health: those with lower levels of offline support reported improvements in their depression scores when receiving online support (Bessière et al., 2008) and higher health perceptions (Rains and Keating, 2011). In a further complication to the interrelated nature of offline and online support, Miyata (2002) found that receiving support online promoted both the receipt and provision of offline support. Consequently, Internet interest, existing social embeddedness, and attitudes about how the Internet affects relationships and society are incorporated into this research.

Concerns about negative influences of social support on health, for example through feelings of dependency or obligation to reciprocate (La Gaipa, 1990; Due et al., 1999, see subsection 2.1.2) may be somewhat ameliorated, as computer-mediated weak tie communication implies less long-term dependency and obligation to reciprocate (Walther and Boyd, 2002). Support may still be inappropriate, for example in forums on anorexia or those glorifying suicide, and anonymity may invite disrespectful behaviour (Wright and Bell, 2003). Particularly for perceived health, Kross et al. (2013) observed slightly negative effects in the context of Facebook usage: more usage led to declines in how people felt during the day, and how satisfied they were with their lives. However, Kross et al.'s (2013) study did not control for offline social support, and other research confirmed positive effects for online social networks (Nabi et al., 2013). In general, empirical research has documented hardly any real harm from social support online (Eysenbach et al., 2004).

2.3.4 RQ_2: Mechanisms of Internet use and health

The different ways in which the Internet may influence perceived health and health service use inspired the second research question about its mechanisms. Bringing together the literature on the effects of Internet use with its influences on health is important for strengthening the theoretical understanding the relationship between Internet use, perceived health and health service use. Moreover, from an analytical point of view, suggesting credible mechanisms belongs to the prerequisites for corroborating the causality of findings (Miles and Huberman, 1994). Consequently, this research will analyse the following RQ in the second results chapter (Chapter 5):

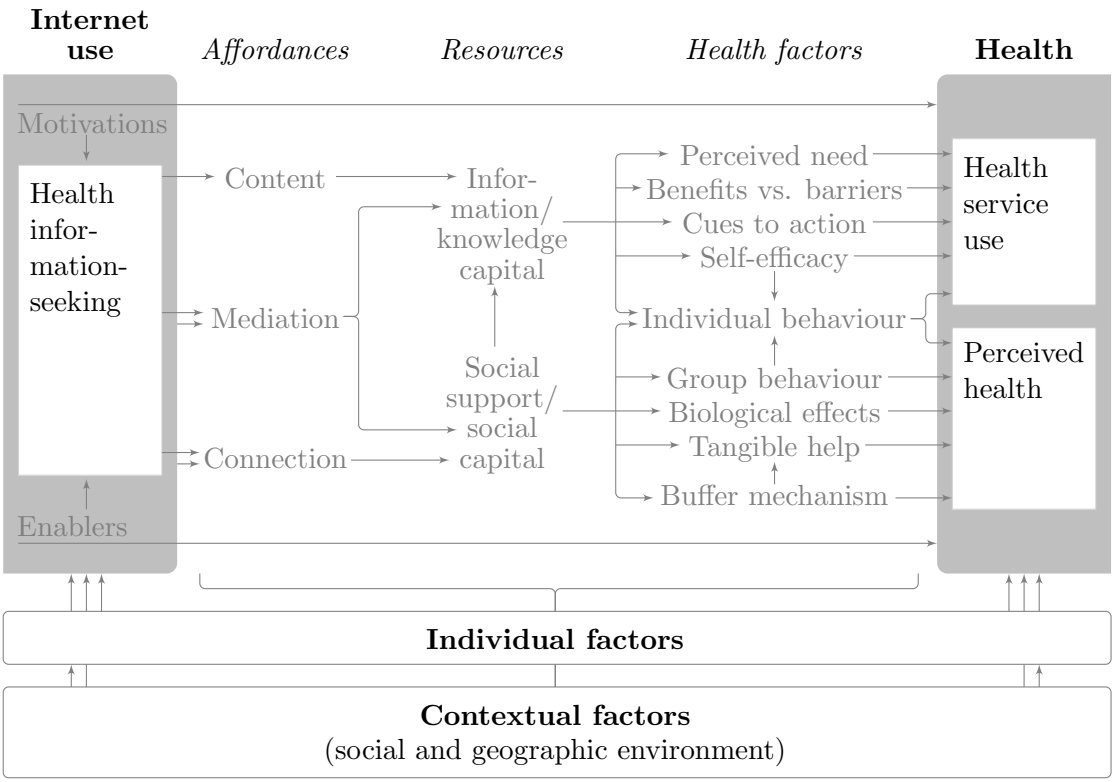
RQ_2 What are the mechanisms by which Internet use affects health service use and perceived health?

As previously discussed, health may also influence Internet use, as implied in Andersen's (1995) feedback loop from health and health service use shown in the analytical framework (Figure 1). While these are not the main focus of this research, they have already been integrated into section 2.2 on motivations (such as health services leading to follow-up questions) or into the enabling factors (such as health-related factors that impede individuals' Internet use and outcomes). The final part of the analytical framework not yet discussed centres around individual and contextual factors, which will be presented in the next and final section of this literature review.

2.4 Individual and contextual factors

The individual's sociodemographic characteristics, as well as the social and geographic environment, have been found to influence Internet use (Papacharissi and Rubin, 2000), its antecedent factors such as motivations and expected gratifications (Cho et al., 2003; Rubin, 2009), technology awareness and desire for information (Straubhaar et al., 2012), the extent to which individuals are able to realise outcomes (Cho et al., 2003; Eynon and Malmberg, 2012), and health more generally (Adler and Newman, 2002). Researching these factors is important due to their influence on health outcomes, the *relationship* between Internet use and health, and as dimensions of the digital divide (van Dijk, 1999; DiMaggio and Hargittai, 2001). In addition, these factors also present a frame for both RQ_1 and RQ_2. Figure 5 shows the updated – and thereby completed – framework, with each part detailed in this section: individual factors in subsection 2.4.1, and contextual factors in subsections 2.4.2 (social environment) and 2.4.3 (geographic context).

FIGURE 5: Analytical framework (4/4) – individual/contextual factors



Individual and contextual influences can be theoretically conceptualised based on the longstanding debate around connecting macro and micro levels of social reality, or the dualism of structure and agency (Turner, 2005; Fuchs, 2008), with Bourdieu (1984, 1986) and Giddens (1984) among the most important thinkers. Both agency and structure are visible in the analytical framework: agency as individuals' free will to pursue actions is represented by motivations and resulting activities (such as health-related Internet use), and structure as the factors that enable and constrain action (Giddens, 1984) included in the enablers, individual and contextual factors. Contemporary researchers generally concur that agency and structure form a dialectical relationship; it is their relative importance that remains controversial (Bottero, 2004; Cockerham, 2005; Fuchs, 2008). This research primarily seeks to understand what their interplay looks like in the context of the Internet and health, in what ways they influence perceived health and health service use, and the theoretical and practical implications for Internet-based health outcomes.

Incorporating these factors also provides the conceptual foundation for both parts of the mixed methods research design, as will be followed up in Chapter 3. In the

quantitative part, individual and contextual factors form the basis for the so-called "constraint variables" in the SM process, as will be explained in subsection 3.2.2, and the control variables for all quantitative analyses to adjust the results for these known influences on Internet use and health (see section 1.2). In the qualitative part, individual and contextual factors are the basis for ensuring breadth in the maximum variation sample in terms of individual and area-based sociodemographic characteristics (see subsection 3.3.1), and the basis for interpreting the interview data while paying attention to its socially constructed nature (see subsection 3.4.4).

Finally, understanding individual and contextual factors also has practical motivations. As Bandura (2004) argues, there exists an ongoing debate about approaches that highlight individual agency, in the sense that individuals should assume full responsibility for their health and exercise control, and structuralist approaches that rely on social, environmental, economic and political factors for improving population health. Across many nations, including the UK, there has been a shift to move responsibility to the individual (Taylor and Bury, 2007) – whereas "in actuality, health promotion needs both approaches, not contentious debates" (Bandura, 2004, p. 159). In addition, interventions may be developed based on certain individual and contextual factors. Although factors such as education and socioeconomic status may not be changed in the short-term (Andersen, 1995), these may be powerful vehicles for shaping health beliefs and behaviours as factors influencing the two health outcomes in this research (Rosenstock, 1966).

2.4.1 Demographics

The main demographic factors described in the context of Internet use and health are related to age, gender, education, socioeconomic status, ethnicity and long-term health conditions, as touched upon in section 2.1. In previous empirical studies, these factors were often included only as control variables (for example Bessière et al., 2008; Berger et al., 2005; Suziedelyte, 2012), or not controlled for at all (Nabi et al., 2013). This subsection will detail all of these factors.

Blank's (2015) "First Law of the Internet" captures **age** differences in Internet use: "everything is related to age, [as] on the Internet young people do more of everything" (p. 3). This relates to use and non-use, intensity of activities, and skills (van Deursen et al., 2011), although the elderly do not form a homogeneous group (van Deursen and Helsper, 2015). While older individuals may have more health problems, research found that younger individuals were more likely to look for health information (Cotten and Gupta, 2004; Powell et al., 2011; Suziedelyte, 2012),

partly due to the connections between health with fitness and lifestyle. Actual and perceived health decline with age (Pampalon et al., 2007), although their correlation becomes weaker, as older people may still perceive their health as good based on their adjusted expectations (Ferraro, 1980; Leinonen et al., 2002; Schneider et al., 2004). Similarly, while older people may need more health services (Dening et al., 1998), middle-aged people in particular use preventative services (Rosenstock, 1966).

For **gender**, differences in Internet use have continuously narrowed over the last years (Dutton and Blank, 2013), although women may be more likely to search for health information online and engage in preventative health behaviours (Atkinson et al., 2009; Bansil et al., 2006; Suziedelyte, 2012; Ybarra and Suman, 2008). Despite similar actual abilities, women also tend to rate their Internet skills lower (Hargittai and Shafer, 2006). In terms of health service use, females see a doctor more frequently, particularly for preventative services (Rosenstock, 1966), which may be explained by women bringing ailments to the doctor's attention more quickly (Hannay, 1980). However, these differences in service use nearly disappear when adjusting for gender-specific conditions (Mustard et al., 1998). For perceived health, women tend to rate their health slightly lower (Fillenbaum, 1979; Pampalon et al., 2007), which may partially be due to women associating psychosocial wellbeing more strongly with their health perceptions than men (Blaxter, 1990; Hansen et al., 2001).

Education remains a strong predictor for Internet use and health (Mirowsky and Ross, 2008). General and health-related Internet use was higher among those with a higher level of education (Dutton and Blank, 2013; Atkinson et al., 2009; Bansil et al., 2006), although those with lower levels of education also used health resources online (Gustafson et al., 1999). However, searching for health information online can be challenging due to low Internet literacy or health literacy (Birru et al., 2004), and skills in both using the technology, and finding and evaluating content were correlated with higher education levels (van Deursen et al., 2011). Education is a key determinant of health in general (Chandola and Jenkins, 2014) and perceived health (Pampalon et al., 2007), partially because education influences health behaviours and beliefs (Lee et al., 2004). In line with this finding, educated people show a better health status, and use health services more frequently only for preventative services (Rosenstock, 1966).

A similar picture arises for **socioeconomic status**, with those from lower socioeconomic groups (and lower income) using the Internet less for health information-seeking (Bansil et al., 2006; Bessière et al., 2008), obtaining fewer benefits online (van Deursen and Helsper, 2015), having lower health perceptions (Pampalon et al.,

2007), and tending to distrust physicians (Bell, 2014). For health service use, early US research found that individuals from higher classes sought treatment more frequently (Koos, 1954; Rosenstock, 1966), although with mixed findings more recently (Bell, 2014; Suziedelyte, 2012). This discrepancy may be due to the difference between health need and health service use: while lower classes showed more actual symptoms, they usually have more difficulties interpreting them (Koos, 1954), feel higher barriers to help-seeking, and display more negative attitudes towards the health system (Ludwig and Gibson, 1969). Moreover, as already touched upon in subsection 2.1.2, health disparities due to differences in socioeconomic status may be explained by different social networks, which encourage or discourage certain health behaviours (Lindström et al., 2000).

Findings for **ethnicity** differ. Some researchers found differences in health information-seeking (Bansil et al., 2006; Suziedelyte, 2012), whereas others did not (Fox and Rainie, 2002). Gustafson et al. (2001) concluded that ethnicity was confounded with socioeconomic status or income, and that differences in health-related Internet use and outcomes stemmed rather from these factors, just as health service use based on online health information were not related to ethnicity when controlling for potential confounders (Suziedelyte, 2012). In general, however, ethnicity may be less applicable in the UK context due to the different historic context of ethnicity in the UK (compared to the US), as has been found by other research on Internet use in the UK context (Blank, 2013).

Finally, those with **long-term health conditions** tend to look for health information more frequently, most likely because health information-seeking increases with the number of health problems, while controlling for other covariates (Rice, 2006; Ayers and Kronenfeld, 2007; Bansil et al., 2006; Bundorf et al., 2006). However, effects on health service use are unclear: some chronically ill individuals may rather use more services based on their higher need – and people with long-term conditions now account for 50% of GP appointments and more than 70% of inpatient bed days (Department of Health, 2012d). Then again, others opted for self-care due to building up experience and expertise, also dependent on past experiences of their illness and health services (Gately et al., 2007; Rogers and Nicolaas, 1998). However, actual effects on health service use through better knowledge about chronic conditions are disputed, as evaluations of the UK's Expert Patient Programme have shown (Taylor and Bury, 2007; Gately et al., 2007). Importantly, while those with long-term health conditions may suffer from poorer health, the "disability paradox" reveals that they may still perceive their health as very good, particularly once

individuals experience enriched meaning to their lives after having adapted to the new normal level (Albrecht and Devlieger, 1999).

As touched upon in section 2.1, these sociodemographic structural factors affect perceived health and health service use in various ways, partly through health behaviours and health beliefs (Adler and Newman, 2002; Rosenstock, 1966; Andersen, 1995). This leads to an additional concept in the area of agency and structure: Bourdieu's (1977) *habitus* is the way through which individuals, largely unconsciously, incorporate structure into their routine practices (Bourdieu, 1977). It is internalised as a consequence of individuals' history and environment, so that "members of the same group are the product of the same objective conditions and share a habitus without realizing that their practices are harmonized beyond what they as individual agents know or wish" (Straubhaar et al., 2012, p. 7). Consequently, individuals from certain sociodemographic backgrounds may differ in their practices of everyday life information-seeking (Savolainen, 1995), interact differently with the health system, or display different health behaviours (Hale, 2013), as how individuals are "treating [their body], caring for it, feeding it, maintaining it [...] reveals the deepest dispositions of the habitus" (Bourdieu, 1984, p. 190). Understanding habitus hence adds an important theoretical element, as it explains the link between individual and contextual features and the rest of the analytical framework.

Habitus operates in relation to specific *fields* (Jenkins, 2001). Bourdieu (1984) refers to the field as a system of social positions in which individual agency may be shaped by specific structural forces (Jenkins, 2001). Fields are therefore defined both by the available capital, and by the agents' struggles to maintain or acquire it (Straubhaar et al., 2012). While Bourdieu does not give clear examples of actual fields, Kleinman's (1978) different domains of medical systems as cultural systems help to conceptualise fields here. For example, Kleinman's (1978) professional domain suggests that a field of relevance to this research might be the medical consultation, where individuals may find themselves in different positions based on different role sets (Merton, 1957) assigned to members of the field. Another possible field may be the private sphere, corresponding to Kleinman's (1978) popular domain, where individuals exist in the context of their social circles and their neighbourhood.

While Giddens' (1984) and Bourdieu's (1986) theories both address the interplay between structure and agency, they are still distinct in their theoretical foundations, the relative importance of structure and agency, their view of individuals as knowledgeable and reflexive agents, and their conceptualisation of resources (Lizardo, 2010; Tucker, 1998). Relevant differences for this research are well reflected in the

concepts of habitus and field. As regards the first, Bourdieu describes habitus as an unconscious way in which structures are embodied in individuals, and although habitus does not determine action, it largely shapes them, with some possible reflection on the part of the individual. Giddens, in contrast, sees individuals as reflexive agents that do certain activities based on their intentions, and views social relations between reflexive agents and the conditions influencing them as more explicit (Tucker, 1998; Fuchs, 2008). With respect to fields, while Giddens talks about time-space relations or locales as contexts constituting interactions, he does not describe the equivalent of a field, which rather encompasses an entire ecosystem of relations not necessarily bounded by time and space. These distinctions between Giddens (1984) and Bourdieu (1986) will be referred to again where applicable in the discussion in section 7.3, while the next two subsections introduce structural factors from the social and geographic environment.

2.4.2 Social environment

The role of the social environment has already been touched upon in multiple instances throughout this chapter. Section 2.1 described how individuals use others to compare their symptoms to, while others also approve or disapprove of certain health behaviours. In addition, those with harmonious relationships with others tended to rate their health as higher despite health problems (Albrecht and Devlieger, 1999). Support was also associated with more physical activity (Poortinga, 2006), suggesting that some socioeconomic differences in health behaviours may be traced back to the social environment (Lindström et al., 2001). Finally, the social environment may provide social support (Wellman and Frank, 2001), and is therefore an integral part of the analytical framework in Figure 5.

The same holds true for Internet use. Section 2.2 described the different user concepts, where the role of the social environment became apparent in the context of the support divide. For non-users, the social environment plays a role for access to the Internet, while for users, it shapes their uses of the technology (Rubin, 2009) and may help them to make sense of content found online (DiMaggio and Hargittai, 2001). Technical help may be equally important as "emotional reinforcement from friends and family, in the forms of both commiseration when things go wrong and positive interest in sharing discoveries when things go right" (DiMaggio and Hargittai, 2001, p. 12). Both would increase motivations for technology use and user competence, which are important predecessors to actual use and building up resources online (Straubhaar et al., 2012).

Section 2.3 introduced the idea of the Internet potentially helping in building up social capital and in exchanging social support. Several researchers found that existing offline social support determines the effects of Internet use: the Internet may increase loneliness for those who already have a wide social circle offline (McKenna and Bargh, 1999), while it had a particularly positive effect on depression for those with low offline social support (Bessière et al., 2008). In support of the interrelated nature of online and offline social circles, Miyata (2002) observed that the receipt of online support also increased the giving and receiving of support offline. A corollary is that social and digital inclusion are closely linked (Helsper, 2008, 2012), in the sense that the socially excluded may be less likely to engage online, while digital engagement might also affect social inclusion as an important influence on perceived health (Schaefer et al., 1981) and health service use, with faster readmissions to hospital due to insufficient social integration (Reblin and Uchino, 2008).

Finally, there is a link between the social and geographic environment, which is both a consequence of social structures and reproduces them (Bottero, 2004). Better social integration[1] in neighbourhoods was in principle associated with lower prevalence of mental health conditions, fewer unhealthy behaviours (Stockdale et al., 2007) and better health (Browning and Cagney, 2002), although the positive effects may depend on the kind of contacts established (Cattell, 2001; Ferlander, 2007). The Internet may contribute to this by fostering how individuals are connected, particularly with respect to their local environment (Hampton and Wellman, 2003). The effects of geographic location will therefore be explored in the next subsection.

2.4.3 Geographic environment

As the last contextual factor, the geographical context may also influence general and health-related Internet use. Those in rural areas use the Internet less frequently in general and for health information (Hale et al., 2010), and health differs with respect to where people live (Kawachi and Berkman, 2003; Macintyre and Ellaway, 2003; Pickett and Pearl, 2001; Tunstall et al., 2004, see section 1.2). This both relates to health service use (Andersen and Davidson, 2007) and health perceptions (Blaxter, 1990; Poortinga, 2006; Soobader and LeClere, 1999). While findings are inconclusive on whether these differences remain significant once individual socioeconomic and health-related factors are controlled for (Kawachi et al., 1999; Pampalon et al., 2007;

[1] Across publications, this relates to social capital, social cohesion and collective efficacy (Browning and Cagney, 2002; Ferlander, 2007; Campbell, 2001), including Putnam's (2000) social capital as civic engagement, which is different from the social capital in the analytical framework.

Browning and Cagney, 2002), place-related influences may operate through various mechanisms on different geographic levels (Curtis and Rees Jones, 1998; Bernard et al., 2007), as briefly outlined in this subsection.

Place has been defined inconsistently in the literature, and is enmeshed with similar concepts such as space, location or locale (Tunstall et al., 2004; Curtis and Rees Jones, 1998). In this research, the definition of geographic contextual factors follows Giddens' (1984) notion that places, or "locales" in his terminology, are the settings of social interactions and represent a structural element critical to understanding individual practice. Place is therefore defined as a "unique system of health-relevant resources and social relationships embedded within geographical borders" (Bernard et al., 2007, p. 1841, drawing on Curtis and Rees Jones, 1998). In that sense, place is similar to Bourdieu's (1984) concept of fields, which, however, is not necessarily geographically bounded (see subsection 2.4.1). "Place" may refer to large or small geographic entities, and factors at all levels may influence health (Soobader and LeClere, 1999; Spielman and Yoo, 2009): neighbourhoods represent the more immediate social contacts, stores, restaurants and health services, whereas factors around labour markets, public services and education may also be determined at higher geographic levels (Bernard et al., 2007).

In this research, geographic place is not only included in the analytical framework and data analyses used to understand the relevance of place for the relationship between the Internet and health. It is also deeply engrained in the area-based approach of the research design, which uses OAs – as small geographic entities with about 300 individuals, as opposed to neighbourhoods with about 5,000 individuals (Power, 2000) – area classifications, urban/rural distinctions and regions within the country in the qualitative and quantitative enquiry (see Chapter 3). Moreover, geography is conceptually reflected in the choice of England for this research. Many US-based models presented in section 2.1 included influences on the accessibility of care, health policy, financing and organisational questions (Andersen, 1995), the structure of the health service delivery (Mechanic, 1975), treatment availability (Mechanic, 1978; Suchman, 1965) and the cultural values of the health system (Kleinman, 1978), just as the empirical literature focused on health insurance coverage or saw its non-inclusion as a limitation (Suziedelyte, 2012; Lee, 2008). However, these factors are less relevant in this UK-focused research due to the relatively uniform access to health services that are free at the point of delivery (see subsection 1.2.2).

On the one hand, compositional features of neighbourhoods (the people living within the area) may influence people through cultural norms, values and social net-

works (Macintyre and Ellaway, 2003). In addition, although poorer areas also have social capital, individuals do not meet better-off individuals and as a consequence are excluded from resource-rich networks (Nast and Blokland, 2014; Ferlander, 2007). Moreover, the social environment in poorer areas may reinforce existing structures and resulting (unhealthy) practices (Campbell, 2001). This shows how the compositional factors relate to different parts of the analytical framework, such as social capital and group behaviour.

On the other hand, contextual features may also play a role: these include environmental hazards, recreation facilities, Internet broadband access, and health provision, for example between urban and rural areas (Macintyre and Ellaway, 2003), partly because geographical proximity to healthcare facilities plays a role for health service use (LaVela et al., 2004). Not only actual health hazards, but merely the perception of the environment in terms of access to leisure facilities, services provided and the social environment in general shaped individuals' perceived health (Poortinga, 2006; Pampalon et al., 2007), although the precise mechanisms (via the social environment or behaviours encouraged in a certain area) remain disputed.

To a certain extent one may choose where to live, but selection of place is also determined by individual resources and social networks (Kawachi and Berkman, 2003), particularly for where one grows up (Dorling et al., 2001). This "reflection of both chances and choices" (Bernard et al., 2007, p. 1840) depicts the link between individual and contextual characteristics. Finally, places are not static, which may be explained through the interplay between agency and structure: the structure of places strongly influence residents' behaviours, who reify that structure through their practices, so that contextual and compositional characteristics change jointly over time (Bernard et al., 2007). Macintyre and Ellaway's (2003) conclusion that "people create places, and places create people" (p. 26) summarises the bidirectional importance of the spatial dimension incorporated in this research.

2.4.4 RQ_3: Individual and contextual influences

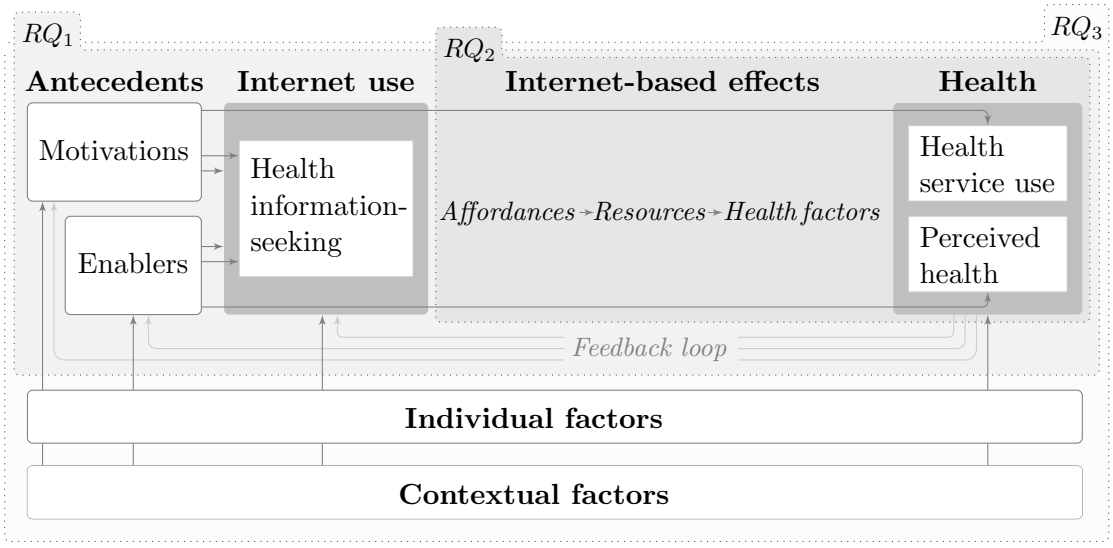
This section presented the importance of individual and contextual factors, and their influence on both Internet use and the two health outcomes in this research. Supported by theoretical grounding around agency and structure, these factors may also be important for the relationship between how Internet use shapes perceived health and health service use, as "health care use is most appropriately viewed as an interplay between agency and structure rather than the outcome of 'expressed need', individual decision-making or 'supply induced' demand" (Rogers and Nicolaas, 1998,

no page). The third and last RQ to be addressed in Chapter 6 of this research will therefore read as follows:

RQ_3 How do individual and contextual factors influence the relationship of Internet use, health service use and perceived health?

These three RQs cover different parts of the analytical framework introduced in Figure 1, as shown in Figure 6. RQ_1 relates to the relationship between different kinds of Internet users, their motivations and enablers, and health outcomes. RQ_2 focuses on *how* these are related, disentangling the mechanisms connecting Internet use, health service use and perceived health. RQ_3 sets the analytical framework in the wider context by incorporating individual and contextual factors shaping the relationship.

Figure 6: Mapping of research questions to analytical framework



Before these RQs are answered in Chapters 4–6, the next chapter introduces the mixed methods methodology, explains how the concepts introduced throughout this chapter are represented in the data, and lays out the way in which the duality of individual and contextual factors from the analytical framework translates into the research methodology by which these RQs are explored.

3 Methodology

Assessing the relationship between Internet use and health presupposes taking a view on how reality and knowledge are constructed (Lincoln and Guba, 1985). Among the plethora of world views, the positivist end rests upon the assumption that there is one generalisable truth about how the Internet influences health – the "one and only truth that is out there waiting to be discovered by objective and value-free inquiry" (Feilzer, 2010, p. 6). At the interpretivist end, many truths about the relationship between the two may exist, making generalisability neither desirable nor possible (Miles and Huberman, 1994; Tashakkori and Teddlie, 1998).

Medical research has traditionally adopted a rather positivist stance, aiming at a high degree of generalisability, assuming that everything can be measured objectively following the rules of the natural sciences and seeing randomised controlled trials as the gold standard in intervention studies (Bowling, 2009; Song et al., 2010). At the same time, health-related social science research has acknowledged the complexity of factors of human influence on health research (Pope and Mays, 1995) and a shift towards patient-centred healthcare (Glogowska, 2011), and hence has frequently viewed health phenomena from an interpretivist angle.

In the "contested methodological arena of health services research" (Barbour, 1999, p. 39), this research acknowledges the right for both philosophies to exist, and adopts the pragmatic paradigm as a compromise between the two poles. This world view regards knowledge as "both constructed *and* based on the reality of the world we experience and live in" (Johnson and Onwuegbuzie, 2004, p. 18, emphasis in original) and gives priority to the choice of the best available method, or combination

of methods, to address the problem at hand. The pragmatic paradigm is especially useful in health research, where problems are usually multidimensional and multidisciplinary, and hence may benefit from a wide spectrum of approaches both for advancing theory and deriving practical recommendations (Ong, 1993). By choosing the pragmatic paradigm, this research enables the findings to be generalised, which is important for analysing how individual voices can be collectively organised to develop practical and policy recommendations on either individual or area level. At the same time, it avoids missing out on the "complexity and messiness of social life" (Feilzer, 2010, p. 14), which is pertinent both to how individuals use the Internet, to how this influences their health service use, and – even more so – to how subjective perceptions of their own health may be shaped.

Stemming from the pragmatic research paradigm, this research employs a mixed methods approach introduced in the next section (3.1), followed by deeper analysis of each of the two parts: section 3.2 focuses on the quantitative data selection, formation and validation, while section 3.3 explains the qualitative sampling and data collection. Both parts are reintegrated in section 3.4, which lays out the analytical strategy used to generate insights from both data sources jointly and separately. Section 3.5 concludes with the most relevant ethical challenges and the chosen countermeasures in this research.

3.1 Mixed methods research design

The mixed methods design employed in this research contains three phases: a quantitative, a qualitative, and a final quantitative phase (see Table 3).

Table 3: Mixed methods research design

Enquiry	**Conceptual research design**		**Data source**
Quantitative	*Large-scale data* Analysis of theory-derived questions for RQ_1–RQ_3	*Large-scale data* Follow-up of emerging interview themes for RQ_1–RQ_3	Secondary data: OxIS (2013), HES (2013), Census (2011)
Qualitative	*43 interviews* Validation/contextualisation of quantitative results for RQ_1–RQ_3		Researcher's own data collection in semi-structured interviews

The combination of both quantitative and qualitative enquiry helps to "best attack a research problem with an arsenal of methods that have nonoverlapping weaknesses in addition to their complementary strengths" (Brewer and Hunter, 1989, p. 17). While the quantitative part provided a large-scale view of the connections between Internet use and health outcomes, it may not have fully explained the relationship (Pope and Mays, 1995) and may have complicated making causal inferences with certainty (Bowling, 2009). To fill this gap, the qualitative strand provided "a concrete, vivid, meaningful flavor that often proves far more convincing to a reader – another researcher, a policymaker, or a practitioner – than pages of summarized numbers" (Miles and Huberman, 1994, p. 1).

In addition to the advantages of triangulation and the complementarity of research findings, this research has benefitted from mixing methods to use findings from one stage for developing the next (Greene et al., 1989). The preliminary analysis of the quantitative data influenced the collection of the qualitative data, thereby ensuring that the interviews would capture not only key concepts identified in the literature, but also initial quantitative insights. The qualitative findings were then used for deeper analysis of the quantitative data, modelling different paths of effects and testing if emerging themes appeared on a larger scale (Creswell and Clark, 2011). Doing so also reduced the risk of being constrained by the "boxes" implied by the quantitative data, which may conceal basic social processes (Silverman, 2006), especially when using secondary data not designed for the research purpose. This design corresponds closely to Mechanic's (1979) ideal study for researching the effects on health and health service use. He envisioned "a broad multivariate design", where researchers would obtain quantitative data, and then "would at least obtain detailed qualitative information from a subsample to help enrich our understanding of the behavioral processes involved" (p. 395). This would create benefits in terms of representativeness and make it possible to use statistical controls based on a wide range of variables, while still capturing individuals' perceptions and understanding the underlying social processes (Mechanic, 1979).

Following the key criteria for distinguishing different kinds of mixed methods designs (Creswell and Clark, 2011; Morgan, 1998), the quantitative and qualitative streams in this "bracketed" design (Greene et al., 1989, p. 263) were addressed sequentially, and were of equal priority in answering RQ_1–RQ_3 as derived from the pragmatic paradigm. The quantitative data formation through SM and its initial analysis took place in May and June 2014 and informed the qualitative interview design. Qualitative data was collected in 43 semi-structured interviews from July

to November 2014, including transcription and initial data analysis, followed by further analysis between December 2014 and February 2015. After incorporating the interview insights, the second part of the quantitative data analysis was conducted from February to December 2015, along with writing up the results.

Both mixed methods strands are also connected in the sampling approach. With OxIS being used in the quantitative part of this research, 31 out of the 43 interviews followed up former OxIS participants in 14 selected OAs, which are areas formed based on the census data for statistical purposes (covered in more detail in subsection 3.2.2). The other 12 out of the 43 interviews were conducted with purposively selected individuals from the same OA, as detailed in subsection 3.3.1. Using the same sample in both strands of the mixed methods design is generally advisable if the intention is to corroborate or compare their findings (Creswell and Clark, 2011). In addition, choosing OxIS participants for the qualitative strand added to the feasibility and efficiency of conducting the qualitative research, and helped to identify characteristics of the qualitative interviewees (such as long-term health conditions) in advance of the interview (Kemper et al., 2003; Onwuegbuzie and Collins, 2007). Most saliently, using OxIS as a base for sampling helped to ensure the breadth of the qualitative sample by building on the features of probability sampling in the qualitative strand, while complementing it with the purposive sampling of additional interviewees where appropriate (Kemper et al., 2003).

With this sampling approach, the parts of the mixed methods design are linked on both conceptual layers of the tiered analytical framework with individual and contextual influences: individuals and OAs. In the quantitative strand, OAs formed the basis of joining the secondary datasets, and were used as the first level of the multi-stage sampling strategy in OxIS, the survey data employed in this research. The qualitative strand then followed up 14 of these OAs used in OxIS, interviewing individuals in these areas. The duality of individual and contextual levels also crosses to the data analysis stage, where the quantitative part analyses the geographic dispersion of the outcome variables (see section 6.3), while the qualitative part employs a coding scheme with influences on both levels.

3.2 Quantitative data formation

In the quantitative strand, this research employs three datasets: the Oxford Internet Surveys (OxIS, 2013) by the Oxford Internet Institute (OII), the English census (2011) by the Office of National Statistics (ONS), and Hospital Episode Statistics

(HES, 2013) by the HSCIC. They are connected through SM to analyse the relationship between Internet use (from OxIS), perceived health (from the census) and health service use (from HES).

This connection is necessary as no single dataset exists to research this relationship. Potentially eligible datasets – such as the British Household Panel Survey (BHPS)/UK Household Longitudinal Survey (UKHLS), the Opinions & Lifestyle Survey (OLS), and the Adult Media Use (AMU) survey by Ofcom – have only limited detail on Internet use and its antecedent factors, or do not include relevant health outcomes (see overview in Appendix VIII, p. 310). However, these datasets are still valuable for external validation of the simulation in subsection 3.2.4.

3.2.1 Selection of datasets

Datasets were purposively selected based on their suitability for answering the RQs. Table 4 summarises all three datasets used in this research. Alternatives to HES were closely evaluated, but rejected due to covering too small a part of the population,[2] focusing only on selected procedures,[3] or not being coded on OAs, and hence not being connectable to the other datasets.[4]

OxIS is a series of biennial surveys that were begun in 2003. They ask a wide range of questions on how people use the Internet, their attitudes, skills and their offline context. As the interviews take place offline using a traditional paper-and-pen method, the sample includes all kinds of users and non-users of the Internet. OxIS consists of repeated cross-sectional surveys, randomly sampling about 2,000 individuals from Great Britain in the multi-stage approach mentioned previously: first OAs and then households within these are randomly sampled, as interviewers are given addresses in an area and ask for the person with the closest birthdate in the household. OxIS also includes a range of communal establishments such as estates for the elderly, so that only those without a permanent home are excluded. To ensure the resulting sample matches the population characteristics, all observations are weighted accordingly (Dutton and Blank, 2013).

The **census** is a decennial count of the entire English and Welsh population, and includes both households and communal establishments. Released on OA level as the lowest level of aggregation, it provides data for gender, age, education, socioeconomic

[2] Clinical Practice Research Database (CPRD), BHPS/UKHLS, OLS

[3] Patient Reported Outcome Measures

[4] General Health Questionnaire, National Wellbeing project, GP Survey, Quality and Outcomes Framework, Public Health Outcomes Framework, BHPS/UKHLS

Table 4: Employed datasets matched to analytical framework



	OxIS	Census	HES
Source	OII	ONS	HSCIC
Method	Survey	Survey	Procedural data
Collection	Offline	Offline	Offline
Sampling	Multi-stage random sample	Complete population	Population who used health services
Geography	England, Wales, Scotland	England, Wales	England
Aggregation	Individual (including OA)	Aggregated (based on OA)	Aggregated (based on OA)
Years/frequency	Biennial (2003–2013)	Decennial (..., 2001, 2011)	Annual (1989–2013)
Individuals England, ≥ 16 years	2013: 2,150 in 260 OAs	2011: 42,989,620 in 171,372 OAs	2013/14: 13,481,805[5] in 171,372 OAs
Internet use	User concepts		
	Online activities		
	Internet centrality		
	Reasons for non-use		
Internet-based effects	Internet outcomes		
	Social consequences		
Health	Health-rel. use barriers		
		Perceived health	
			Health service use
Antecedent factors	Motivations		
	Enablers		
Individual factors	Age	Age	Age
	Gender	Gender	Gender
	NS-SeC	NS-SeC	
	Education	Education	
	Long-term condition	Long-term condition	
	Life stage		
Contextual factors	Social environment		
	Geographic environment	Geographic environment	Geographic environment

Note: Appendix IV (p. 295) provides details on all constructs. Linked concepts are used to connect datasets (see Appendix V, p. 302). Appendix VIII (p. 310) compares available constructs to the UKHLS/BHPS, AMU and the OLS.

[5] Refers to hospital *admissions* of individuals ≥ 16 years between April 2013 and March 2014

status (operationalised by the National Statistics Socioeconomic Classification, NS-SeC) and long-term health conditions for every individual living in the OA. Moreover, with "How is your health in general?", everyone rated their health status on a scale from 1 to 5. Although the data is aggregated on an OA level, cross-sections are also published: for example, educational level can be broken down by age (Census, 2011a), while perceived health can be segmented by age, gender, long-term health condition, education and NS-SeC (Census, 2011b,c,d).

Finally, **HES** data contains inpatient hospital stays of the population of England.[6] The dataset released for this research is aggregated on OAs like the census, and includes the number of unique individuals who went to hospital each month[7], their primary diagnosis, and gender/age groups. As described in Appendix VI (p. 306), the outcome variable for this research was formed by dividing the number of unique individuals in hospital each month by the number of individuals in the same age/gender category in the OA, so that the resulting outcome concept ranges from 0 to 12 (if everyone in the OA in the age/gender group was hospitalised at least once every month). For selected analyses in section 5.3, the outcome was restricted to admissions for stigmatised or mental health conditions only (see Appendix VI, p. 307) to make it possible to compare results to the literature that focused on these outcomes (for example Berger et al., 2005). As HES data is automatically created by procedural data, it does not contain any self-reported patient measures.

With these three datasets, this research uses data from 2011 (the census) and 2013 (OxIS and HES). The difference of two years rests upon the fact that the most recent census is from 2011, while OxIS saw irregularities in the OA information in 2011. Moreover, using OxIS data from 2013 instead of 2011 meant a higher chance for the qualitative interviews that contact data had not changed. This approach is in line with previous SM research, which reported differences between three and six years between the sample and the census (Smith et al., 2009; Tomintz et al., 2008; Anderson and Speed, 2009). Geronimus and Bound's (1998) US-based research found that due to the relative stability of area-based measures over time, there are no problems associated with using earlier census data.

While longitudinal analysis with 2011/2013 and 2001/2003 (with a difference of 10 years between each dataset) was explored, this additional analysis was not pursued further. First, OxIS (2003) has some known data quality differences in

[6] Outpatient data was requested, but was discarded due to data quality problems.

[7] Number of actual visits were also requested, but the item "Count of events" seems to include more than visits, ranging up to several thousand events for one individual per month.

comparison to later years and is therefore left out for most of the longitudinal analyses in Dutton and Blank (2011, 2013). Moreover, several variables important for data analysis and the SM process, such as long-term condition and NS-SeC, were not collected at all. An alternative approach – using the same survey dataset (OxIS 2013) while combining it with the outcome variables for both years (census for 2011 and 2001; HES for 2013 and 2003) – was not viable due to the relative stability of these outcomes on the small-area level over time (Geronimus and Bound, 1998).

Out of the various geographies for which data is available, this research uses only England (everyone above the age of 16), hence covering a total of 42,989,620 individuals. In addition to the substantial motivations for choosing England explained in the introduction, the choice was also dictated by practical reasons. First of all, due to the organisational structure of the NHS (it is divided by country), HES data is not available for Scotland or Wales in the same way. Secondly, in England and Scotland, OAs – the connecting element between the datasets detailed below – differ in their definition and socioeconomic homogeneity requirements (ONS, 2013b), so that including Scotland would introduce an undesirable bias into the analysis.

3.2.2 Combination through spatial microsimulation

The three datasets are combined through SM, a small-area estimation technique to combine individual survey data with aggregated, area-based data to create a large-scale dataset estimating individuals' attributes (Ballas et al., 2005). For this, a representative sample survey with rich information on the aspect under consideration (Internet use in OxIS) and a second dataset with limited data items, but high geographic disaggregation (the census), were combined (Cassells et al., 2013). After the SM process, HES data was added for every small area, which enables Internet use, perceived health and health service use to be analysed jointly (Deetjen and Powell, 2016a). Using SM – as opposed to just appending the respective outcome values to OxIS – has advantages in terms of randomising the error for point estimates, and increases the dataset's practical value due to its applicability for small area analyses across England, as detailed in subsection 3.4.3, as shown in the Isle of Wight case study (subsection 6.3.3), and as corroborated in Appendix XI (p. 332).

The unit of analysis is the individual (as opposed to the household); this unit was chosen for substantive reasons (see Chapter 2) and because all three datasets were available for the individual level, even if distributed in aggregated form. All datasets generally include households and communal establishments, which is important as the scope of the datasets in SM should match as closely as possible (Cassells

et al., 2013). From a health perspective, those in closed care homes or staying in hospital frequently were less likely to be selected, and might have generally refused to be interviewed (Bowling, 2009). Indeed, the OxIS population with a long-term condition (16%) is lower than the census suggests (21%), which may also be due to the slightly different question wording (see Appendix IV, p. 300).

The basis for combining the three datasets are OAs, which exist in OxIS for every respondent as the first stage of the sampling process, and constitute the aggregation level for both census and HES data. OAs are areas consisting of around 300 individuals and are formed based on geographic and socioeconomic characteristics to achieve homogeneity within groups (ONS, 2013b). However, they are arbitrarily determined spatial units as their boundaries are not visible in the real world (Dietz, 2002). The approximately equal population size implies that OAs vary in geographic size: they may relate to one street in central London, or cover a whole village or a collection of rural hamlets in sparsely populated settings elsewhere. To ensure equal sizes and socioeconomic homogeneity, OAs are a result of each census rather than an input, and hence may change. This is relevant for this research as OxIS 2013 uses 2001 OA definitions, while the census uses 2011 definitions (and HES includes both). However, from 2001 to 2011, this change was only about 2.6%, with a large share of OAs being either split or merged (ONS, 2013b). For those 260 OAs in OxIS 2013, only one OA was split between 2001 and 2011, but as the area characteristics (in terms of urban/rural and area classification) were homogeneous after the split, this did not have implications for the SM process.

In short, the simulation works as follows: for every OA – for which the aggregate characteristics are known from the census – the algorithm determines which of the individuals from the survey (OxIS) need to be chosen to best replicate who lives in the respective OA, with HES data being tied to the resulting dataset based on the spatial location. As a spatial model, this provides a simulated dataset of all individuals in a given geographic area with their Internet use, Internet-based effects, motivations, enablers and social environment (from OxIS), self-rated perceived health (from the census), and health service use (from HES).

Before going into detail, it is useful to discuss the criteria used to select these individuals: the so-called "constraint variables". Constraints are variables shared between the datasets involved in the SM process which allow for the modelling of the spatial distribution of the survey data. The general rule is that the more constraints, and the more constraints for variable interactions, the better (Williamson, 2013), although there is a danger of model over-specification (Lovelace and Dumont, 2015).

Typically, researchers in this area use four to five constraints (O'Donoghue et al., 2013). The constraints must fulfil three conditions: first, they must be correlated with the variables taken to the small-area level (Internet access/use); second, they should be correlated with the variable to be analysed at the small-area level (perceived health/health service use); and third, they need to exist in both the census and the survey, sharing the same definition and categories (Anderson, 2013).

Fulfilling these conditions (see Appendix V, p. 302), this research uses gender, age, education, socioeconomic status (NS-SeC), long-term condition, and locality as constraints. Ethnicity was excluded, as it had no significant effect on either Internet use or health, which is in line with the previous literature (see subsection 2.4.1). As SM allows using cross-tabulations (Edwards and Tanton, 2013), age and education were included as interrelated constraints. This is useful as the population's formal education levels have increased in recent decades (Census, 2011a). It is worth highlighting that OxIS as a random sample covers nearly all potential category combinations, but not necessarily all of them. However, some combinations may not even exist in the population, and covering all potential combinations is no prerequisite for SM (as long as the sample is random), as the census does not consist of a joint cross-tabulation of all constraints either.

The inclusion of locality is a special feature of this SM model. While other SM models have incorporated regions for selecting individuals (Ballas et al., 2005; Anderson, 2013), OxIS – unusually – includes the OA of each respondent, so that aspects captured in the ONS area classification (socioeconomic, ethnic, work-related and urban/rural features) are employed as a cross-tabulation with the nine regions in England (see Appendix V, p. 305). Incorporating locality is important as both (health-related) Internet use (Hale et al., 2010) and health (Riva et al., 2009) differ by location, and individuals are assigned area-based values from the census and HES for the analysis. In addition, as discussed in section 2.4, while areas may consist of similar individuals, they may have different contextual features, so that geodemographic factors such as the OA area classification (ONS, 2014b) should be included in the SM model (Birkin and Clarke, 2012). Finally, factoring in locality also responds to the need "to incorporate greater 'local' specificity" (Whitworth, 2013, p. 29), for example through subsample selection, as is done in this research.

Finally, the order of the constraints matters as well, as variables entered later into the model produce a better fit (Anderson, 2009). A multiple regression with the proposed constraints on Internet access and use helped to determine the order of constraint variables simultaneously, as in the stepwise regression in Anderson

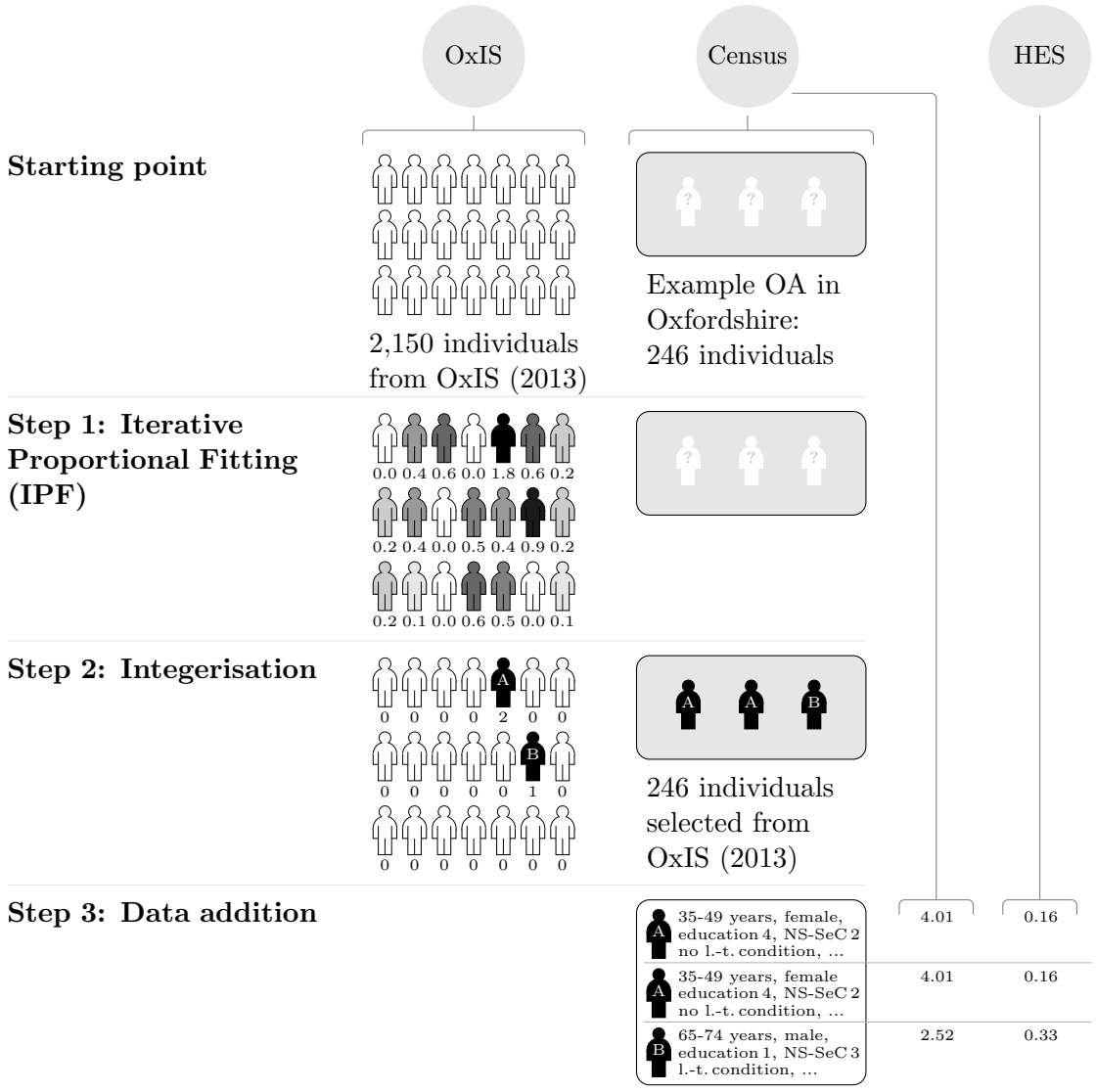
(2009), while separate evaluations of Goodman-Kruskal's γ with significance based on the χ^2-test were used for the aggregated census variables. One complication was that the use of multiple target variables produced slightly different orders (see Appendix V, p. 302). As a compromise, the variables were used in the model in the following order of increasing relevance: locality, socioeconomic status, gender, long-term health condition, education and age.

With these constraints, the Internet-related characteristics of individuals in each OA are now determined through "reweighting an existing microdata sample [...] so that it would fit small area population statistics tables" (Ballas et al., 2006, p. 68). Among the plethora of approaches that in principle result in similar datasets (Whitworth, 2013; Rahman et al., 2010), this research employs iterative proportional fitting (IPF): observations are weighted for each OA based on the constraints and may be integerised subsequently to analyse whole individuals (Ballas and Anderson, 2013; Lovelace and Ballas, 2013). Deming and Stephan (1940) and Fienberg (1970) derived the mathematical basis of IPF, Birkin and Clarke (1988) and Wong (1992) described its use in geographic analyses, and Lovelace and Dumont (2015) detailed its implementation in the statistical package R. Appendix IX (p. 316) contains the mathematical foundations and their translation to R, as used in this research.

Figure 7 shows the SM process for one example OA in Oxfordshire. Step 1 consists of determining the weights for each OxIS respondent to replicate the OAs in terms of the constraints. If there are 246 individuals above the age of 16 (out of 300 individuals in total) in the example OA in Oxfordshire, 115 men and 131 women, 65/65/39/13/64 individuals from education levels 0–4 and so on, then these numbers will be closely replicated by taking 0.21 of individual 1, 2.13 of individual 2, 1.0 of individual 3... up to 0.7 of individual 2,150 (the numbers are examples only). Individuals from a similar OA in the South East region are given higher priority, although this locality constraint is released if no suitable individuals are available (see Appendix V, p. 305). The "iterative" in IPF reveals that these weights are determined through convergence to a solution. To achieve this convergence, the research used 10 iterations as a compromise between results and computational speed (Ballas et al., 2005; Anderson, 2007a, as cited in Anderson, 2007b).

As 0.21 or 2.13 represent unnatural fractions of individuals, Step 2 of the process is integerisation (see Figure 7). In the context of this research, one might leave out this step with the intention of reducing the additional error of the SM model (Anderson, 2009, 2013). However, this research uses integerisation since it gives better interpretability of the simulation, for example with policymakers, and reduces

Figure 7: Schematic summary of spatial microsimulation



Result: Simulated dataset with 42,989,620 individuals (everyone in England $\geq$ 16 years) in 171,372 OAs, with rich data on Internet use and related constructs (from OxIS), perceived health (from the census) and health service use (from HES data)

the size to about one seventh compared to the non-integerised version. From the different integerisation methods available, this research uses the truncate-replicate-sample (TRS) algorithm by Lovelace and Ballas (2013) to obtain whole individuals. TRS combines the probabilistic element of integerisation, which ensures that each individual with a weight larger than 1 is selected at least once in the final dataset

and generally outperforms alternative deterministic approaches regardless of which validation measure is evaluated (Lovelace and Ballas, 2013). Some individuals who are particularly representative of the OA may be selected more than once, as in the case of Person A in the example OA in Figure 7. An exception is section 6.3, which contains an analysis for one specific local authority district using both integerised and non-integerised data for comparison.

The interim result after integerisation is a simulated dataset with everyone in England over 16 years and their Internet use. Still missing is Step 3 – adding the health data from the census and HES. Crucial here is not only to use the average for all 246 individuals in the example OA, but rather the average for the specific gender, age, education, long-term condition and socioeconomic status (for perceived health from the census), and gender and age (for health service use from HES). Creating individual data from aggregate data is known as "ecological inference"; it includes approaches such as least-square regression (Goodman, 1953), Duncan and Davis' (1953) methods of bounds, and more recently various combined strategies (King, 1997; King et al., 2004). In this case, the outcome variables consist of the mean of three cross-tabulations: perceived health by age, gender and long-term condition (Census, 2011b); perceived health by education (Census, 2011c); and perceived health by NS-SeC (Census, 2011d). Forming the outcome variables in this way strikes a compromise between the correctness of the estimated values, and the complexity and comprehensibility of the results. As some tables (Census, 2011b,c) are published only with perceived health recoded to 1–3 – (very) bad health; fair health; (very) good health – due to statistical disclosure control, the first and the latter category were split again based on the proportion of bad versus very bad (and good versus very good) provided per OA in the census (2011d). Appendix VI (p. 306) shows a detailed conceptual overview of how the outcomes were created.

After repeating this procedure for all 171,372 OAs, the result of the SM procedure is a simulated dataset with 42,989,620 individuals, which now allows for the joint examination of associations between Internet use, perceived health and health service use. Some caveats for analysing this simulated dataset – due to its size, its non-independent simulated individuals, and its multi-level structure of OAs and individuals – will be detailed in the analytical strategy (section 3.4). In addition, based on the probabilistic nature of the integerisation approach, there is not one single dataset alone. However, evaluating differences for 10 of these final datasets indicated that their divergence is very small, as will be shown in the next subsection (alongside other ways of validating the generated output).

3.2.3 Internal validation

"Statistical model validation is a key (arguably 'the' key) part of the process of model building" (Edwards and Tanton, 2013, p. 250) and a prerequisite before data analysis (Ballas et al., 2005). Internal validation verifies the SM results using the datasets included in the simulation, and helps to ensure that the SM process itself has been successful (Edwards et al., 2011; Lovelace and Dumont, 2015). To begin with, the implications of validation are more easily understood by comparing the actual census proportions to the simulated proportions (Ballas et al., 2005). Table 5 takes one arbitrarily selected region in England – South East based on the example OA in Oxford – and shows that differences between both are relatively small. While ethnicity was rejected as a constraint, deviations between actual and simulated values were relatively small too (Asian: 4.9%/5.2%; Black: 1.4%/2.3%; White: 91.8%/90.9%; Other: 1.8%/1.5%).

Table 5: Simulated vs. actual values for South East region

Age	*simulated*	*actual*	**Gender**	*simulated*	*actual*
16–24	13.8%	13.9%	Male	48.3%	48.6%
25–34	15.2%	15.3%	Female	51.7%	51.4%
35–49	26.7%	26.8%			
50–64	22.7%	22.9%	**Long-term health condition**		
65–74	11.0%	10.9%	Yes	17.8%	17.9%
75–84	7.3%	7.2%	No	82.2%	82.1%
≥ 85	3.1%	3.1%			
Education			**NS-SeC (socioeconomic status)**		
No qualifications	19.3%	19.1%	Class 1: Professional occ.	35.2%	35.3%
Level 1 (primary)	18.5%	18.7%	Class 2: Intermediate occ.	24.1%	24.2%
Level 2 (secondary)	19.3%	19.5%	Class 3: Routine/manual	29.3%	29.0%
Level 3 (further)	13.1%	12.8%	Long-term unemployed	4.1%	4.1%
Level 4 (higher)	29.7%	29.9%	Full-time students	7.3%	7.5%

Source: Actual values taken from Census (2011a,b,f,g)

For a more formal validation, Appendix VII (p. 308) shows the values for Pearson's R, the Chi-square test, the total absolute error (TAE), the standardised absolute error (SAE), and the (normalised) root mean square error (RMSE and NRMSE) as well as deviations larger than 5% and five individuals across the different constraint variables. There is no generally accepted goodness-of-fit measure as all of them have their benefits and drawbacks (Voas and Williamson, 2001), so that this research reports a number of different measures and their interpretation below.

For all measures, it is important to keep in mind that the error stems from the explicit choice to integerise the simulated dataset (see subsection 3.2.2). This does not apply for the analyses in subsection 6.3.3, where the Isle of Wight as a specific local authority district was analysed based on non-integerised data as well.

Pearson's R is used as a first indicator for goodness-of-fit to evaluate how well the simulated totals for each of the constraints correlate with the actual totals. The average value of Pearson's R obtained based on 10 simulated datasets across all constraints is 0.984 (age: 0.991; education: 0.975; long-term condition[8]: 0.994; gender: 0.949; NS-SeC: 0.977; area (in the wide sense): 1.000). It is not uncommon to reach values near 1 in constraint evaluation, with 0.9 as the recommended threshold for acceptance (Lovelace and Dumont, 2015). Similarly, the Chi-square test is used to examine whether differences between the actual and the simulated data are due to chance only. It had a p-value of 1 across all constraints except for area in the narrow sense ($p = 0.93$) after removal of all areas in which any of the cell values were smaller than 5. In line with the Pearson's R values, the result confirms that the differences are very small and based on chance only.

Other measures to evaluate the SM process are those that sum up the difference between simulated and actual totals across all areas and constraints. The TAE, defined as the sum of all absolute deviations, is hardly meaningful due to the size of the modelled area. Instead, Voas and Williamson's (2001) SAE evaluates actual versus simulated totals relative to the total population (individuals $\geq$ 16 years in England). It ranges from 0 (perfect fit) to 2 (no fit) and was 0.17 (17%) in this research. While there is no uniformly accepted threshold, the SAE should be below 20% in about 80% of the areas (Clarke and Madden, 2001, as cited in Smith et al., 2009). This is the case here: 79.7% of the OAs have a SAE of lower than 20%.

The RMSE and the NRMSE further penalise larger deviations, while keeping the measures comparable through standardising for the number of zones and – in the case of the NRMSE – the range of observations. They were 30.73 and 1.31% respectively, which is relatively small compared to other models (Lovelace et al., 2015). Given that perfect fit is not the aim, another way to evaluate fit is to analyse the number of OAs in which the simulated and actual values deviate by more than 5% (Lovelace and Ballas, 2013). This is the case for 27.87% of the cells, in line with other models

[8] About 14% of OAs were excluded from this calculation, as the total sum (long-term health condition: yes and no) differed considerably from the number of people in the OA. As the simulated numbers need to be formed based on the total number of people in the OA, the actual and simulated values rightly differ, and hence were excluded from the validation calculation

(Lovelace and Ballas, 2013; Lovelace et al., 2015). However, as constraint variables have rather different category sizes, measuring absolute deviation may be more useful. The proportion of OAs in which the simulated constraint value deviates more than five individuals from the actual values is 3.74%. These four measures confirm that there are relatively few large deviations.

A final internal validation check analysed how many individuals were replicated in the areas where they are actually located, which was made possible by the recording in OxIS of the individual's OA. On average, the simulated dataset replicates about 71% of individuals in their original areas. It is natural that this number does not reach 100% due to the probabilistic nature of how individuals are selected by the TRS algorithm, particularly for relatively homogeneous OAs in terms of the constraints, as opposed to extreme areas where only few individuals may be suitable for cloning within the locality constraint (Birkin and Clarke, 2012). However, aggregating several consecutive simulations, the proportion of individuals replicated in their areas asymptotically approaches 100%, and reaches 97% based on 10 SM datasets (see Appendix VII, p. 309). This implies that no individuals are systematically excluded from being replicated in their areas.

3.2.4 External validation

External validation involves the comparison of SM output with other datasets to evaluate whether the simulated dataset is a good representation of the actual world (Edwards and Tanton, 2013). This subsection shows that this is the case in this research, provided that traditionally collected surveys properly reflect the real world.

External validation is often difficult to perform due to a lack of suitable datasets (Whitworth, 2013), but in this case there were three datasets available (see Appendix VIII, p. 310): the Internet access module of the Opinions & Lifestyle Survey (OLS, 2013) by the ONS, the Adult Media Use (AMU) survey by the UK communications regulator Ofcom (2013), and the UK Household Longitudinal Study (UKHLS, 2013), which replaced the British Household Panel Survey (BHPS). All of these are not nearly as fine-grained as OxIS in terms of individuals' Internet use, but allow for a comparison of the overall levels of Internet use, as well as perceived health and health service use, as they were investigated – along with general Internet use – in the same survey in the OLS and the UKHLS. Table 6 shows that the simulated values for general and health-related Internet use fall within the range of values in the OLS, AMU and UKHLS and their 95% confidence intervals. Appendix VIII (p. 313) provides a detailed breakdown across different regions of England.

Table 6: External validation – comparison of proportions of individuals

	Source	**Simulation**	**External validation datasets**		
Category	*OxIS*		*OLS*	*AMU*	*UKHLS*
Internet users	79.1	78.8	85.1	77.5	83.3
95% confidence interval	[77.3; 80.8]	[78.8; 78.8]	[83.6; 86.4]	[75.1; 79.7]	[82.9; 83.7]
Individuals in hospital	–	3.4	–	–	4.7
95% confidence interval		[3.4; 3.4]			[4.5; 5.0]
Health information-seekers,					
at least monthly	32.7	31.1	44.0	31.0	–
95% confidence interval	[30.4; 35.1]	[31.1; 31.1]	[41.9; 46.1]	[28.2; 33.9]	
at least weekly	13.2	11.3	–	10.0	–
95% confidence interval	[11.5; 15.0]	[11.3; 11.3]		[8.3; 12.0]	

Note: Internet users and individuals in hospital are measured in % of the population; health information-seekers in % of Internet users.

It is worth noting that there are some differences in the question wording for each of these datasets (see Appendix VIII, p. 312). With respect to the proportion of Internet users, OxIS asked whether individuals "personally use the Internet on whatever device at home, work, school, college or elsewhere", whereas AMU listed all devices (personal computer, laptop, netbook or alternative device), but restricted the question to the home. In contrast to both of these datasets, the OLS and the UKHLS asked for general use without specifying devices or locations. This may explain the higher proportion of Internet users in the OLS and UKHLS, which goes beyond the limits of the confidence intervals for OxIS, the simulated data and AMU.

Health service use is higher for the UKHLS than for the simulated dataset, as the UKHLS does not differentiate between inpatient and outpatient visits, whereas HES only relates to inpatient services in this research. However, as Appendix VIII (p. 315) shows, health service use is lower for older individuals in the UKHLS. A potential explanation is that the UKHLS only measures whether health service use was mentioned in response to the question "Which of these services have you used over the last 12 months?" As older individuals tend to include age-related problems in their health assessments (Schneider et al., 2004; Kivinen et al., 1998; Leinonen et al., 2002), they may also view health service use as a normal activity.

Finally, for the proportion of health information-seekers, OxIS and AMU asked about the frequency of online health information-seeking (with "at least monthly" as one of the answers), whereas the OLS referred to having looked for health-related content online in the last 3 months. It is natural that the latter is slightly higher.

Additionally, the OLS and the UKHLS feature the question of how individuals rate their health status. This enables the relationship between Internet use and perceived health to be validated using the same dataset in Table 7. Based on the slightly different scales and question wording of the census, OLS and UKHLS, this comparative analysis employs perceived health on a scale from 1 to 3 (see Appendix VIII, p. 314).

Table 7: External validation – perceived health

	Source	**Simulation**	**External validation datasets**		
Categories	*OxIS*		*OLS*	*AMU*	*UKHLS*
User	–	2.79	2.81	–	2.80
95% confidence interval		[2.79; 2.79]	[2.78; 2.83]		[2.79; 2.81]
Non-user	–	2.55	2.36	–	2.39
95% confidence interval		[2.55; 2.55]	[2.28; 2.44]		[2.36; 2.41]

Note: Perceived health is measured on a scale from 1 to 3, as the census and the OLS employed a scale from 1 to 5 with 3 as the neutral option, while the UKHLS used 2 as the neutral option.

Both surveys corroborate the evidence from the simulated dataset that non-users have lower perceived health. The relationship also holds true when the numbers are split by age and gender, though with smaller differences in the means of users versus non-users for the lower age groups (see Appendix VIII, p. 314). The overall lower health for non-users in the OLS and the UKHLS can be attributed to the different proportions of respondents with long-term conditions. While the census shows that about 21% of the English population has a long-term condition, this was the case for 35% in both the OLS and the UKHLS, in contrast to OxIS (which the simulated dataset is based upon), according to which only 16% had a long-term condition. In addition, it is important to keep in mind the different scale for the UKHLS with 2 as the neutral option (5: Excellent; 4: Very good; 3: Good; 2: Fair; 1: Poor), which may have elicited different responses.

Another form of validation is qualitative data (Pope and Mays, 2006), testing the RQs on the same individuals used in OxIS by asking them directly about their health-related Internet use and its effects. Stemming from the pragmatic research paradigm, however, the qualitative part is far more than a simple validation of the quantitative data: it represents a different way of answering the RQs, and enriches and contextualises the findings beyond the quantitative numbers.

3.3 Qualitative data collection

For the qualitative part of this research, 43 individuals from all over England were interviewed face-to-face for about 50 minutes on average, with a minimum of 25 minutes and a maximum of 105 minutes. The interviews revolved around how they use the Internet in general and for health-related purposes, their motivations and enablers for use, and how all of these factors influenced decisions to see a doctor and the way they perceived their health (see interview guide in Appendix XVII, p. 349).

The 43 interviews were conducted in three major waves: the first 10 interviewees in three areas (Maidstone, King's Lynn & West Norfolk, and South Norfolk) were interviewed between 22 and 30 July 2014. The second wave (Birmingham, Amber Valley, Middlesbrough, Gateshead, Cornwall, East Dorset, and the Isle of Wight), with 21 interviewees, was conducted between 18 August and 25 September 2014. The final wave (Ealing, Hammersmith & Fulham, Wiltshire, and Newcastle-under-Lyme), with 12 interviewees, took place between 12 and 26 November 2014. Conducting the interviews with a staggered approach created room for "self-reflexivity about subjective values, biases, and inclinations" (Tracy, 2010, p. 848), and made it possible to refine questions to pursue emerging aspects, as Pope et al. (2006) recommend as an ongoing process during data collection. The fieldwork breaks also served for transcribing interviews, documenting ideas and observations, and completing an initial round of coding (Saldaña, 2012).

3.3.1 Sampling approach

The 43 interviewees were sampled from the 260 OAs in OxIS (2013), with 31 of them being OxIS 2013 respondents themselves. The qualitative sample therefore extended the multi-stage sampling approach employed in the quantitative strand, and took advantage of the probability sampling from OxIS (Kemper et al., 2003). This enabled maximum variation in terms of sociodemographic characteristics (Sandelowski, 1995b; Coyne, 1997), and allowed for the incorporation of area-based features in the analysis. In addition, using OxIS participants also provided knowledge about interviewee characteristics in advance of the interviews, and opened up valuable ways to move between the quantitative and qualitative data at the analysis stage (see section 3.4). Additional purposive sampling by approaching individuals in public spaces and snowball sampling based on OxIS participants improved the coverage of under-represented groups and provided access to particularly information-rich cases (Sandelowski, 1995b; Pitney and Parker, 2009; Suri, 2011). From a practical point of

view, the area-based interviewee selection meant that several face-to-face interviews could take place within smaller geographic areas.

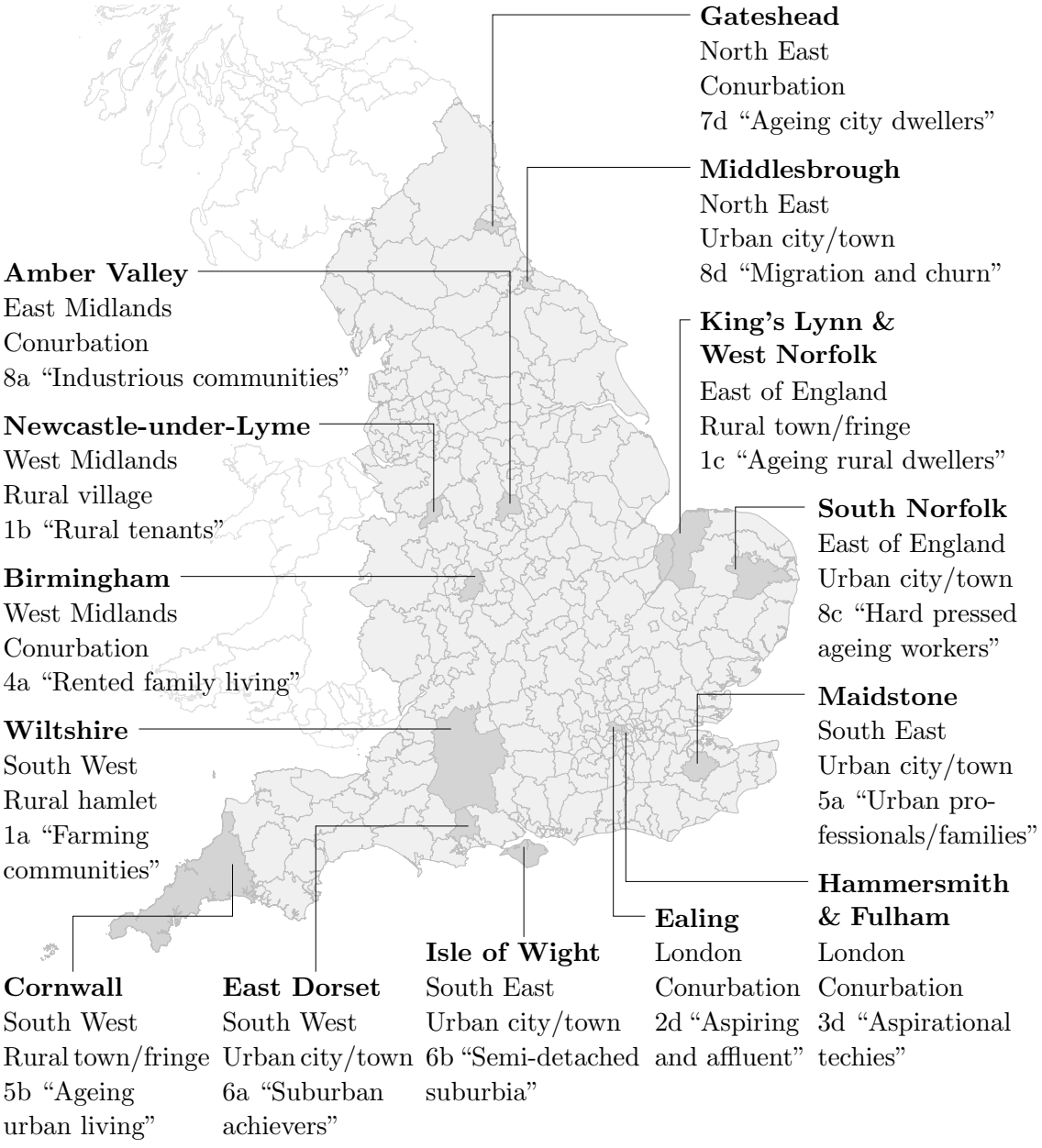
The area selection logic consisted of the four main steps shown in Table 8. The first step was long-listing all OxIS OAs with at least 12 respondents, and requesting their contact details from the company in charge of data collection for OxIS (2013). When the results of the data request came back, 8 OAs were excluded due to having fewer than 9 OxIS respondents who provided a telephone number for follow-up interviews. The remaining 33 OAs were stratified according to three non-mutually exclusive criteria to enhance diversity: OA classification (demographic, household composition, housing, socioeconomic, and employment characteristics (ONS, 2014b)); urban/rural locality; and region. Additionally, preference was given to OAs with at least one interviewee with a long-term health condition to ensure richness of qualitative data (Marshall, 1996).

Table 8: Area selection logic

	June 2014		*July 2014*		*July–November 2014*
	Long list for data request	Exclusion of missing data	Stratification of remaining OAs		**Final decision for each OA**
Criteria	All OxIS OAs with $\geq$ 12 individuals ($\geq$ 18 years) selected	OAs with $\leq$ 9 individuals with contact data excluded	OA classification, locality, and region evaluated	OAs with $\geq$ 1 long-term health condition preferred	Calls 1–2 weeks in advance; $\geq$1 OxIS interviewee for OA to be selected
Aim	Feasibility	Feasibility	Area diversity	Richness of data	Feasibility
Result	41 OAs (561 individuals) longlisted	Of all 41 OAs, 8 excluded	Of the remaining 33 OAs, 14 shortlisted, 19 on backup list		14 OAs selected (10 from shortlist, 4 from backup list)

The final step was contacting all prospective interviewees in the OA via telephone. This was an iterative process during the fieldwork, as calling more than one week in advance turned out to be impracticable for scheduling the interview. Whenever at least one previous OxIS participant was successfully recruited, the OA was selected. This resulted in 14 OAs being selected, as shown in Figure 8 and Appendix XIII (p. 340).

FIGURE 8: Overview of selected areas



SOURCE: ONS rural/urban classification, area classification (Census, 2011j,k).

Within the 14 selected OAs, all OxIS participants with a telephone number were contacted; those with health problems were approached first to increase the chance of arranging a meeting at a convenient time. On the telephone, the researcher provided the OxIS participant with a short introduction (as a "researcher of the University of Oxford"), asked to confirm the interviewee's name, explained they

were called based on their participation in an Oxford survey on Internet use in the previous year, provided a short introduction to the study and asked whether the interviewee would be available for a 45-minute follow-up interview for compensation of 10 GBP in cash or as a donation to charity. The spiel was slightly adapted for Internet non-users, reducing the emphasis on the Internet to mitigate non-response due to unfamiliarity with the technology. Beginning in the second wave, the purpose of the interview as part of a doctoral research project was included, depending on the interviewee's initial reaction, as this information generally increased willingness to participate. This process allowed 31 OxIS participants from 2013 to be recruited.

Additional interviewees were purposively sampled in two ways within the OA: either by approaching individuals in public spaces, or by snowball sampling based on asking the OxIS interviewee who else would be good to talk to in the area. This helped to recruit under-represented groups (Sandelowski, 1995b), such as younger people and males. Younger people were harder to reach due to either having moved away for educational or personal reasons, or having provided a mobile number only which they had ceased to use. Males were already under-represented in OxIS (40%, unweighted), and additionally showed lower interest in participating in the interviews. Through the purposive sampling approach, 12 further interviewees (8 males, 4 females, including 3 in the age groups of 18–24 and 25–34) were recruited, all of whom were relatively frequent online health information-seekers and represented particularly information-rich cases (Pitney and Parker, 2009; Suri, 2011).

As mentioned above, all interviewees were offered 10 GBP compensation to encourage participation (Babbie, 2009). In the first wave of interviews, this was offered in cash only. The additional option of donating the money to charity was added for all subsequent recruitment calls to provide a more universal incentive for individuals from all socioeconomic backgrounds, as motivations for participation may differ based on personal circumstances (Rubin and Rubin, 2004). The interviewees could choose the specific charity they wanted to give the money to, which had the benefit that the choice of (often health-related) charities led to further conversations relevant to the Internet and healthcare. As a result of this research, donations were made to charities for bowel cancer, Parkinson's disease, stroke, rare chromosome disorders, multiple sclerosis (MS) and Ebola. Where interviewees did not specify a recipient, the money was donated to dementia research.

The ratio of successful interviews over number of individuals reached was 29%, with the number of individuals approached being higher due to incorrect telephone numbers. Unfortunately, some individuals dropped out due to having developed

diseases that prevented them from being interviewed (dementia, cancer). However, non-response may be less problematic in qualitative research, as long as the sample is diverse across characteristics deemed important in the analysis (Sandelowski, 1995b). For the sociodemographic determinants of Internet use and health (see sections 1.2 and 2.4), the sample consists of a diverse group, as shown in Table 9.

Table 9: Interviewee demographic characteristics

Gender	*absolute*	*relative*
Male	18	42%
Female	25	58%
Age		
16–17	*(excluded)*	
18–24	3	7%
25–34	6	14%
35–49	9	21%
50–64	10	23%
65–74	7	16%
75–84	5	12%
≥ 85	3	7%

Education	*absolute*	*relative*
No qualifications	7	16%
Level 1 (primary)	12	28%
Level 2 (secondary)	6	14%
Level 3 (further)	6	14%
Level 4 (higher)	12	28%
NS-SeC (socioeconomic status)		
Class 1: Professional occupations	14	33%
Class 2: Intermediate occupations	8	19%
Class 3: Routine/manual	15	35%
Long-term unemployed	4	9%
Full-time students	2	5%

Although the qualitative sample reflects the structure of the population of England relatively well, the primary aim is to have some individuals in each category of these variables to ensure a breadth of different views (Marshall, 1996). This rests upon the idea that a sample may be "statistically nonrepresentative, but informationally representative" (Sandelowski, 1995b, p. 181), as qualitative research is concerned with obtaining diverse accounts rather than providing a representative picture. In fact, it is "often the infrequent gem that puts other data into perspective, that becomes the central key to understanding the data" (Morse, 1995, p. 148), and that helps to shed a different light on the quantitative data (Pope and Mays, 2006). In that sense, achieving maximum variation may be more important than saturation, which is a problematic concept as there may always be new aspects to discover with every unique interviewee (Wray et al., 2007; O'Reilly and Parker, 2012).

In terms of diversity, interviewees covered the whole continuum from perfectly healthy to seriously limited in their daily activities, and made use of the full scale from 1 to 5 when rating their health. The conditions discussed in the interviews covered all of the MCDs (European Commission, 2015), either with the interviewees themselves being affected, someone they provided care for (children, parents or

partners), or friends and acquaintances. Beyond the MCDs, pregnancy, in-vitro fertilisation (IVF), men's health, alcoholism, smoking, dieting/weight, eczema and rare chromosome disorders were covered (see Appendix XIV, p. 341). With respect to Internet use, 20 interviewees were next-generation users, 11 first-generation users, 2 ex-users, and 10 never-users. Out of the 12 non-users, 7 used the Internet by proxy.

Appendix XV (p. 342) provides an overview of all 43 interviewees and their main sociodemographic, health-related and Internet-related characteristics. All interviewees were given pseudonyms starting with the first letter of their local government district (the first two letters in Maidstone and Middlesbrough), for example Anthony from Amber Valley or Brenda from Birmingham. The names reflect gender and ethnicity, and age as far as possible. The interviewee overview also specifies which individuals were recruited purposively, and by whom they were recommended, as these interviewees may be more similar to each other (Pitney and Parker, 2009; Bottero, 2004) and have a generally wider social network (Cohen and Arieli, 2011).

3.3.2 Interview design

Nearly all interviews were conducted face-to-face due to easier rapport-building (O'Connor et al., 2008), particularly for sensitive health topics. Online interviewing with asynchronous communication and some level of anonymity might have been suitable as well (James and Busher, 2009), but was impractical due to the intention to include non-users, the chosen area-based approach, and the use of telephone numbers as contact information for former OxIS participants. Two interviews were conducted over the phone at the interviewees' request (see Appendix XV, p. 342). However, several interviewees emphasised they would not want to talk on the phone:

> *The phone rings, and you want to do things, [...] you got a thing off the top of your head. Plus, when you sit and you're talking, it's sitting and talking, it's listening, it's different. I prefer this way, and I'll be honest with you: you probably won't get the response that you're getting now, the facts would be sort of hazy [...] and as soon as you may hear a question that sounds a bit obtrusive from that person on the phone, you go.*
>
> **– Ian, 42**

The interviewees were allowed to propose a place to meet for the interview (Yeo et al., 2013). As a result, 35 of the face-to-face interviews took place in the interviewee's home, with the others taking place at the interviewee's work place or in public houses/cafés. For all meetings at individuals' homes, a second researcher was notified before and after the meeting for safety purposes (Arksey and Knight, 1999).

From a methodological point of view, meeting the interviewee at home entailed the benefit of seeing the devices used to access the Internet, while some interviewees also felt more protected within the confines of their own home (Berg and Lune, 2014):

> *I'm not very good at talking to people. I think that's because I am in my own house, you see, I'm fine in my own house.*
>
> – **Susan, 58**

The interviews were semi-structured, as structure helped to ensure comparability across interviews, to reference back to the analytical framework, and to connect the strands of the mixed methods design (Bryman, 2008; Creswell and Clark, 2011), while enabling flexibility and probing (Gray, 2009). It also allowed the language to be adapted to the interviewee (Berg and Lune, 2014), and enabled interviewees to "share in the work and the fun of discovery, often guiding the questioning in channels of their own choosing" (Rubin and Rubin, 2004, p. 14) as an important feature of the patient-centric research design. The interview guide consisted of closed and open-ended questions, focusing on the latter to encourage communication and to ensure that the participants' full views on complex health topics were gathered (Arksey and Knight, 1999; Bowling, 2009). Appendix XVII (p. 349) includes the final interview guide. It contains the main questions, divided into separate stages, which were memorised to let the conversation flow more naturally (Gill et al., 2008).

Before the interview, an introduction (highlighting commonalities with the interviewee where possible) and an informal chat not directly related to the research helped to get to know each other, ease the tension and build up initial trust (Rubin and Rubin, 2004). Getting the relationship off to a good start is important as interview data may be seen as socially constructed in the interchange of the researcher and the interviewee (Creswell and Miller, 2000; Bryman, 2008), as will be detailed in subsection 3.4.4. Rapport-building is an important basis for open discussion, and contributes to validity in semi-structured interviews (Arksey and Knight, 1999). However, the right level of support may be one that makes interviewees comfortable to share their thoughts, but still maintains a professional distance so that the interviewer avoids leading the interviewees' accounts (Seidman, 2013).

The actual interview began with a short explanation setting out the research purpose, obtaining informed consent (see Appendix XVI, p. 346), asking for permission for recording, and offering the chance to ask questions (Gray, 2009). As ice-breaker questions, the initial part of the interview revolved around the area as a non-personal topic (Arksey and Knight, 1999), which also gave first insights into

the social and area-based factors of the individual. The interview then proceeded by breaking down the RQs into the different parts of a puzzle, thereby presenting the interviewee with less complex questions one at a time (Rubin and Rubin, 2004).

The main part started with the interviewees' use of the Internet, followed by a discussion of health-related aspects of Internet use, which all interviewees answered with a focus on health information. Interviewees were then asked to recall specific situations in which they had a health-related question and to describe them in detail. Jointly with the interviewer's probes about their motivations, considerations, strategies and effects on health and health service use, these provided vividness, richness, nuance, depth and detail to the RQs (Rubin and Rubin, 2004). The structure of this part was similar for users and non-users, though with a focus on alternative sources of information and proxy use for the latter. At the next stage, more open questions encouraged talking about effects of the Internet on health beyond health information. Towards the end of the conversation, interviewees were directly asked about how the Internet affected their perceived health and health service use, thereby asking about the topic of interest in various ways, from broader to increasingly more focused questions (Rubin and Rubin, 2004).

Following the final open question on other aspects that the interviewee considered important (Gray, 2009), interviewees received a debrief along the lines of the RQs, which inspired several of them to add further interesting remarks about health and the Internet. The interview ended with obtaining or (in the case of OxIS participants) verifying the personal demographic characteristics, giving interviewees the chance to talk about topics that they considered of relevance, thanking the interviewee for the contributions, explaining the next steps and how to hear about the results (Gray, 2009), and some further friendly chat, which often provided further information relevant to the research (Rubin and Rubin, 2004; Bryman, 2008).

The interview guide was derived from the analytical framework, complemented in the initial quantitative analysis, pre-tested with two individuals, and refined between the three interview waves. Concepts emerging as important in the initial quantitative analysis were trust in doctors, attitudes to learning and technology, and self-efficacy. Based on the pre-test, some questions were reworded to avoid leading interviewees towards certain responses (Yeo et al., 2013), and to ensure questions were worded in a "warmly human and cooly nonjudgmental" way to create an atmosphere of openness and empathy (Berg and Lune, 2014, p. 122). Based on the initial considerations and pre-testing, the interview guide worked relatively well from the beginning, with only minor wording changes and a more detailed research debrief being incorporated

between the interview waves. Several interviewees stated that they were prompted to think of specific instances they had not thought about before. This reflects the breaking down of the separation between the offline and the online worlds (Orgad, 2009), with the Internet increasingly becoming an integral part of everyday life:

> *I haven't thought about this so much for years, so I enjoyed that. Whenever you said something, I was "oh yeah, that's this...". Because we just take [the Internet] so much for granted our lives now, don't we?*
>
> – **Brian, 54**

Beyond the general framework of the interview guide, the interviews were slightly adapted for the specific individual during the conversation, as interviewees had different conceptions of what constituted the Internet, especially older and less frequent users of it. For example, Susan referred to her Internet use as *"being on Facebook, but not on the Internet"*, while Ethan consistently talked about *"asking Safari for an opinion"*. In addition, as outlined in Chapter 2, the Internet's boundaries are becoming more fluid through the use of mobiles and the Internet of Things, so the interviews explored the interviewee's language to ensure a common understanding (Yeo et al., 2013; Berg and Lune, 2014).

To "keep the door open to continue the discussion or ask additional questions" (Rubin and Rubin, 2004, p. 121), interviewees had the opt-in possibility to receive a three-page summary of the research results (see Appendix XIX, p. 358). These member checks were not only a way to feed back the results to the participants and thank them again for their contribution through a hand-written note or personalised email, thereby continuing the level of rapport built up in the interview (Seidman, 2013), but also a method to check that their different perspectives had been correctly incorporated (Seale, 1999) as an important facet of ensuring quality in qualitative research (Lincoln and Guba, 1985; Creswell and Miller, 2000). Rather than validating a single truth, feeding back the overall findings from all interviews "provides space for additional data, reflection, and complexity" to ensure meaningful coherence of the research insights (Tracy, 2010, p. 848). Appendix XIX (p. 357) contains the interviewee responses to the research summary (p. 358).

3.4 Analytical strategy

The analytical approach followed the logic of the research design: an initial quantitative analysis, then qualitative data collection and analysis, followed by a final quantitative analysis. As the points of interface between the strands of the mixed

methods design were at data analysis or before, the analysis utilised a common frame to connect the datasets (Creswell and Clark, 2011). This common frame stems from the analytical framework elaborated in the literature review, and has been refined across each phase of the empirical investigation by "connecting from the analysis of one set of data to the collection of a second" (Creswell and Clark, 2011, p. 66).

The initial quantitative analysis in advance of the qualitative fieldwork started out with various uni-/bivariate data checks, as data quality is perfect neither in census data nor in medical data more generally (Connelly et al., 2016; Bowling, 2009). Afterwards, an initial multiple regression on perceived health helped to identify particularly relevant constructs (HES was not yet available at this stage; see Appendix XII, p. 339). The aim of this initial sift through the data was to establish the principal dynamics in the data to inform the design of the qualitative interview guide (see Appendix XVII, p. 349), which generated more detailed insights for the RQs under consideration.

3.4.1 Interpretation of interview insights

All interviews were captured with a voice recorder, starting after the interviewee had signed the consent form and was aware of being recorded (British Sociological Association (BSA), 2003). To extract meaning from the data, all recordings were transcribed as an ongoing process during and between the three data collection waves. Verbal ticks ("er", "um") were not transcribed due to their low analytical value (Arksey and Knight, 1999). Colloquial phrases ("ain't", "gonna", "wanna") have been transcribed, but transformed into standard English in the results chapters. Pauses, laughing and fillers such as "you know" were taken out of the quotes without indication to improve readability (Rubin and Rubin, 2004). For topics that were clearly unrelated to the research, only keywords were noted in the transcripts to be able to go back to the recording if necessary (McLellan et al., 2003).

In addition, note-taking during and after the interview about the individual, the technology they used and the area they lived in helped to go "beneath the surface to explore issues that are assumed, implicit, and have become part of participants' common sense" (Tracy, 2010, p. 843) and to integrate tacit knowledge into the thick descriptions of the qualitative insights (Creswell and Miller, 2000). After the interview, the notes were also complemented with a short review of the data in the light of the RQs to develop preliminary answers and refine the RQs throughout the analysis (Silverman, 2013), but also "as a healthy corrective for built-in blind spots" of the researcher (Miles and Huberman, 1994, p. 50).

Analysing the transcripts made it possible to refine these preliminary insights. All interviews were coded using a top-level coding scheme derived from the analytical framework in Chapter 2, which allowed for some direction in the analysis to avoid "line-by-line approaches to coding [...] without any sense of the whole, or any orienting gestalt to the interview data to be managed, or any sense of analytic direction" (Sandelowski, 1995a, pp. 373–374). Lower-level codes emerged from the data to ensure that unexpected themes were not missed (Silverman, 2006), and were obtained by going through the data multiple times, iteratively combining and splitting categories to reconnect them to patterns and storylines (Dey, 2005; Saldaña, 2012). The final coding scheme is included in Appendix XVIII (p. 352).

In addition to the value of the qualitative data in understanding the relationships and mechanisms between Internet use and health, emerging codes from the interview then formed the basis for the final quantitative analysis. While with secondary data, there is no possibility of generating new data even if this is desirable (Freese, 2009), this final data analysis helped in the development of more advanced models with interaction effects and SEMs. These techniques may be superior, particularly for analysing the complex pathways in which health may be affected (Johnson and Wolinsky, 1993), even though SEM only helps to confirm insights when properly backed up theory and further insights (Chin and Todd, 1995; Bollen and Pearl, 2013), for example from qualitative research. Based on the interview insights, further recommendations for variables to be added to the secondary data in the future are provided in subsection 7.5.4.

3.4.2 Development of typologies

The qualitative data also prompted new ways of thinking about the quantitative data, as often observed and desired in mixed methods studies (Pope and Mays, 2006). For example, the qualitative research indicated the existence of various types of orientations towards health information-seeking, with implications for perceived health and health service use. Typologies are a useful way of describing segmentation in the social world (Spencer et al., 2013), can be a useful tool for building theory (Doty and Glick, 1994), and help for developing practical recommendations (Zubin, 1938), even though the boundaries between types may not be impermeable in a complex social reality. Based on the qualitative observations and the theoretical foundations on motivations and enablers of (health-related) Internet use, the key characteristics of these types were deduced; these enabled the qualitatively inspired typology to be replicated for the quantitative data.

The quantitative clustering was done using hierarchical clustering, as this would to some extent enable the number of clusters to emerge from the data – as opposed to k-means clustering, for example (Everitt et al., 2011). The number of types is usually difficult to determine, as at the one extreme everybody is a unique type, while at the other extreme, the types may be too coarse (Shalizi, 2009). In the research here, the qualitative data provided a rough idea of six types, which could then be balanced with the number of potential clusters in the quantitative dendrogram. The underlying distance measure was Ward's linkage with squared Euclidean distances, since it provided better results than other hierarchical methods in general (Ferreira and Hitchcock, 2009), and specifically for OxIS 2013 (Dutton and Blank, 2013).

While there is no gold standard for evaluating clustering solutions (Jain et al., 1999), quantitative approaches include measuring distances within and between clusters, and entropy as a measure of the distribution of individuals in the dendrogram (Zhao and Karypis, 2004). Another beneficial feature when using the same individuals in both parts of the mixed methods design is the ability to validate the quantitatively replicated typology by checking which individuals fall into which type. In the few cases (6 out of 31 interviewed OxIS participants) where the two diverged, the qualitative data was given precedence due to being more current and containing fewer errors in data collection. Based on the pragmatic paradigm, this divergence will be explored in more detail since "quantitative methods that identify potential outliers seek to deal with them by ensuring that they are not the result of errors [whereas] the qualitative tradition focuses on anomalies – or deviant cases – and uses their potential to illuminate by exception" (Barbour, 1999, p. 41).

The qualitative data also serves to bring the typology to life, with Chapter 4 introducing a representative for each type. While multivocality is important for the credibility of qualitative research (Tracy, 2010), and is adhered to throughout the rest of the results presentation, the key to a typology is how different characteristics occur in the same person, making within-case displays of higher analytic value (Miles and Huberman, 1994; Ayres et al., 2003). To avoid the dilemma that "typologists usually achieve parsimony by providing elegant descriptions of their typologies and glossing over the complex processes" (Doty and Glick, 1994, p. 230), the typology of orientations towards health information in Chapter 4 is not just an end in itself, but is used where applicable to the mechanisms between Internet use and health in Chapter 5, and for analysing individual and contextual factors in Chapter 6.

3.4.3 Analysis of large-scale simulated data

In principle, simulated data can be analysed with standard approaches such as multiple (logistic/multinomial) regression and SEM, with weighted least squares estimators for orderable categorical data (Muthén, 1984). Both approaches serve to answer the RQs of this research: RQ_1 is first explored with regressions to understand the general associations between Internet use and the two health outcomes in this research (Wilkinson et al., 1996), and then with SEMs to be able to incorporate causally informed variables such as "Did you ever find anything online that helped improve your health?" (QC.22d in OxIS, 2013), thereby modelling the pathways by which they are related (Bowen and Guo, 2011; Gefen et al., 2000). RQ_2 about the *mechanisms* between Internet use and perceived health and health service use by definition lends itself towards the use of SEMs due to their usefulness in confirming structural pathways for complex phenomena (Bowen and Guo, 2011). Finally, RQ_3 about the influence of individual and contextual factors rather uses regressions for the analysis, as these allow for better replicability of the findings with other datasets and are more amenable to meta-analyses (Gefen et al., 2000).

The fit of these models can be evaluated with conventional goodness-of-fit measures such as the Adjusted R^2. However, based on how the outcome variables were derived from the aggregate data, the R^2 values are inflated, as a large share of the variance is explained by the constraints (see Appendix XI, p. 334). For SEMs, a comparative fit index (CFI) and Tucker-Lewis index (TLI) of > 0.90 and a root mean square error of approximation (RMSEA) < 0.05 is desirable (Browne and Cudeck, 1992; Rigdon, 1996). As with general statistical techniques (Wilkinson et al., 1996), the analyses assume normally distributed residuals (homoscedasticity) without distortions by nonlinearity, influential outliers or collinearity, as measured by both the largest condition index (CI) and the variance inflation factor (VIF). To reduce heteroscedasticity, values for perceived health were taken to the power of 3, whereas health service use was transformed by taking the cube root. Further transformations were applied to selected independent variables to improve linearity where necessary (see Appendix X, p. 328).

In addition to these generic principles, "new methods sometimes require new ways of effectively reporting results" (Anderson et al., 2001, p. 377). For SM, p-values as indicators of significance that the sample results are also true at population-level are less useful in the context of a population dataset (even if the dataset was just simulated). P-values in general have been widely criticised (Carver, 1993; Kirk, 1996; Anderson et al., 2001), and always need to be put into the context in which

they are used (Wasserstein and Lazar, 2016). In a dataset with several million observations, the p-values will nearly always be significant based on the relationship between power, effect sizes and sample size (Cohen, 1992), independently of whether or not the associations are meaningful (Kirk, 1996; Sullivan and Feinn, 2012). One approach may be to adjust the required significance level downward as the sample size increases (Greene, 2003), but this practice is not common, and lacks standards for what constitutes an acceptable significance levels at different sample sizes (Connelly et al., 2016). Consequently, following Anderson et al. (2001) and Carver (1993), the quantitative analyses in this research do not report p-values, but rather focus on the standardised effect sizes (β coefficients) and the odds ratios (OR), which can be interpreted by comparison with other variables in the model. The use of effect sizes only also supports that all results are reported in full, rather than only a subset of significant findings (Garamszegi, 2006). In addition, as Kirk (1996) has proposed, what counts is practical significance: researchers need to judge whether the results are trivial or not, for example based on the size of effects and further insights from the qualitative inquiry. While not using p-values is a corollary of the nature of the SM process, it is also in the spirit of the pragmatic paradigm, which – unlike (post-)positivist paradigms – does not necessarily rely on accepting or rejecting hypotheses to answer the RQs.

The simulated dataset has a multi-level structure with OAs as the first level and individuals as the second due to the clustered sampling approach in OxIS (Hox, 1998; Goldstein, 1999). In addition, as the 2,150 OxIS individuals are replicated multiple times in the SM dataset, the individuals themselves are not independent of one another and may be seen as a further multi-level structure similar to repeated measurements in longitudinal studies (Baltagi, 2008; Goldstein, 1999). There are two major ways of dealing with clustered data: multi-level models in which the variance is split between OAs and for individuals within them; and clustered standard errors (Rabe-Hesketh and Skrondal, 2012; Primo et al., 2007). In this case, the latter may be more appropriate for the majority of analyses in this research, as the outcome variables on the OA level are not really measured at the between level: perceived health and health service use are broken down by constraint (see subsection 3.2.2), so that for perceived health, there are up to 700 possible values[9] per OA (each, on average, has only about 300 people). The second option would therefore be more appropriate, but adjusting standard errors without using p-values

[9] 700 categories = 7 (age) · 2 (gender) · 2 (long-term condition) · 5 (education) · 5 (NS-SeC)

has no effect on the results. Consequently, the data is analysed as a population dataset with a non-hierarchical structure. As OA boundaries are statistical units, whereas area effects may rather occur within environmentally visible units such as neighbourhoods (Dietz, 2002; Spielman and Yoo, 2009), it may be acceptable to use a traditional single-level model without additional complexity created by multi-level structures for analysis and presentation (Goldstein, 1999). At the same time, the non-independence of observations as one of the underlying assumptions of statistical techniques such as regressions is acknowledged as a limitation in this research (Baltagi, 2008; Wilkinson et al., 1996). Wilkinson et al. (1996) suggests that with non-independent observations, researchers should take into account that their true sample may in fact be smaller than assumed. However, with 2,150 participants in OxIS, the sample size for the analyses presented in this research can still be considered acceptable (Babbie, 2009).

Finally, it is worth discussing the benefits of using the SM approach, as opposed to simply appending the adjusted health outcome values to OxIS. First, including more than the 260 OAs contained in OxIS (2013) makes it possible to incorporate the values for perceived health and health service use from all 171,372 OAs. This helped to randomise the error for the different outcomes values for each OxIS individual: the box-and-whisker plots for 20 selected OxIS respondents in Appendix XI (p. 332) show that some are close to the mean of the assigned outcome values, whereas others diverge substantially from it. The analytical gold standard would be if census and HES data were not aggregated, but could be matched one-to-one to OxIS data; however, this is not possible without appropriate consent for good ethical and privacy-related reasons. Secondly, and related to the first point, as some quality problems with the OA coding in OxIS were noted during the qualitative fieldwork, SM – which selects individuals purely on the basis of the OA characteristics (OA classification, rural/urban locality and region) rather than the OA itself – at least minimises a false reliance on individuals actually being in a certain OA. Thirdly, from a practical point of view, the SM-created estimates for all over England are more useful when deriving practical implications for specific areas within England (see section 6.3), which would not be possible by just using OxIS.

However, due to the drawbacks of the SM approach mentioned above, Chapter 4 replicates the results for the OxIS dataset with directly appended values for perceived health and health service use for two specific analyses in sections 4.3 and 4.5, and thereby allows for the evaluation of p-values and the construction of confidence intervals. In addition, Appendix XI (p. 334) reports all further regressions

for the OxIS dataset. These results are reported without adjustment for the risk of obtaining results by chance due to multiple comparisons through Rice's (1989) suggested Bonferroni correction, as it only examines individual p-values rather than the number of significant statistical tests (Moran, 2003; Garamszegi, 2006), entails a lower probability of finding significant results in later analyses (Moran, 2003; Cohen, 1990), and lacks clear criteria of when it should be used (Cabin and Mitchell, 2000; Moran, 2003). In support of this decision, existing research in peer-reviewed journals also only sporadically corrects for the higher probability of committing type II errors (Cabin and Mitchell, 2000). Juxtaposing the SM-based and OxIS-based results shows that, despite some differences in the size of the coefficients, and various relationships not being significant based on the sample of 2,150 individuals, the general relationships are not greatly changed. This supports the general validity of the SM approach, as will be further explored in the next subsection.

3.4.4 Validity and reliability

Key concerns in research are reliability and validity as "yardsticks against which we judge our relative success or failure in measuring things" (Babbie, 2009, p. 150). These concepts differ for qualitative and quantitative inquiries, but are concerned with the quality of the data, and with the results and interpretations for both (Creswell and Clark, 2011). In contrast to quantitative research, qualitative inquiry focuses on aspects around credibility and trustworthiness to assess the rigour of the research (Lincoln and Guba, 1985). This subsection discusses how deriving valid conclusions from the overall mixed methods research design is ensured, and thereby extends the external validation results in subsection 3.2.4, which already showed that the dataset derived from SM is an acceptable approximation of the real world.

Internal validity mainly refers to whether causal conclusions can be drawn (Gray, 2009), for which three conditions must be fulfilled, following Hume's logic (Miles and Huberman, 1994): the relationship must constantly hold true; there must be a plausible mechanism; and there must be temporal precedence of cause and effect. The first condition is addressed through analysing the relationship across contexts – from the simulated dataset to the interviews – bearing in mind that, in accordance with the pragmatic paradigm, there is no "single truth". For the second condition, qualitative data serves to explore the mechanisms, with SEM helping to confirm different paths quantitatively. Even though SEMs do not prove causality and rely on strong theoretical foundations (Bowen and Guo, 2011), it makes causal inferences more plausible (Bollen and Pearl, 2013). The temporal condition cannot easily

be fulfilled without longitudinal data, but can be improved by incorporating the feedback loop (see section 2.1) into the qualitative and quantitative analyses.

For SM, there are challenges of ecological and atomistic fallacies (Goldstein, 1999; Freedman, 1999). Making individual inferences based on aggregate data implies a certain homogeneity of health conditions (Edwards and Clarke, 2013). Consequently, this research takes OAs instead of modelling "a city like London or large area like Yorkshire" (Edwards and Clarke, 2013, p. 70), with the socioeconomic homogeneity of OAs accounted for in their design as an important prerequisite. Compositional effects, such as OAs that featured certain population groups overproportionally (for example, a deprived area with lower health due to the known social determinants of health), are taken into account through simulating the dataset based on the total counts of all constraints per OA. Moreover, assigning everyone the average for the gender, age, education, NS-SeC and long-term condition group (see Appendix VI, p. 306) made it possible to make reasonable approximations of the individual measures, as intended in ecological inference (King, 1997). The constraints used for SM, which are confounders to the relationship between Internet use and health outcomes by definition, must be controlled for in all quantitative analyses, as required for observational data in general (Winship and Morgan, 1999). Finally, mapping some results in Chapter 6 is complicated by the modifiable areal unit problem (MAUP), which captures the fact that spatial representations depend on the level of geographic aggregation: for example, local authority districts or OAs (Dark and Bram, 2007).

Another challenge is data quality, partially due to the use of probabilistic linkage methods (Lahiri and Larsen, 2005), but also resulting from the secondary data employed in this research (Freese, 2009; Connelly et al., 2016). For releases at the small-area level, some census data records have been deliberately swapped to disguise individuals or households with unusual characteristics (ONS, 2011b). OxIS also faces some data quality problems, although the interviews served to corroborate its data quality. For HES inpatient data, the number of unique hospitalised individuals from an OA was 99 in 47,343 instances. As neither the HES data dictionary nor the HSCIC could explain this code, these rows were deleted. The remaining 13.5 million hospital stays corresponded closely to the HSCIC's (2015) official number of 15.5 million in 2013/14. Most of this difference (1.7 million) is explained by excluding those below the age of 16 in this research; 0.1 million were in English hospitals without being residents; and 0.2 million either went to hospital twice within a month (which would only count once in the HES data released for this research), or could not be associated with any OA, for example due to their identification document

not being registered in the hospital. Outpatient data showed even more inexplicable irregularities and could not be matched to the HES official statistics, and hence was not included in this research.

Another challenge is construct validity. Secondary data entails the corollary that data has been collected for different purposes (Freese, 2009), leaving gaps in what can be measured, and making it more difficult for constructs to measure the concepts they are supposed to represent (Gray, 2009). Construct validity is improved through careful item selection, as detailed in Appendix IV (p. 295). Some constructs are composed of several items, many of them reaching Cronbach's α values above 0.8 or 0.9 (see Appendix IV, p. 295). The minimum acceptance criterion was 0.6 in a few rare instances, with 0.7 as the preferred threshold (Singh, 2007).

Moreover, it would also be valuable to measure the effects of Internet use on primary care outcomes. Secondary care data may not be the best source for drawing conclusions about general health service use, particularly due to the gatekeeping effect of GPs in England (see subsection 1.2.2), and as inpatient stays may instead be associated with more serious conditions. However, due to the non-availability of large-scale primary care data and the data quality problems with secondary outpatient data, these health outcomes will be explored only qualitatively. Inpatient data therefore acts as a surrogate for general health service use in the quantitative analyses – with limitations that will further be discussed in subsection 7.5.3.

Reliability in terms of getting consistent results when replicating the inquiry (Babbie, 2009) would be enhanced by repeated data collection. To some extent, the qualitative data responds to the re-test aspect: in case of deviances from the sociodemographic characteristics in OxIS, this research reports those obtained in the interviews. As touched upon in subsection 3.4.2, while both qualitative and quantitative data are important sources of information for answering the RQs, the qualitative data is more current, contains fewer errors in data collection, and does not face the limitations of simulated datasets. Strategies to improve dependability (as the qualitative counterpart of reliability) included asking interviewees to get in touch if they had any further thoughts, and member checks as the "most crucial technique for establishing credibility" (Lincoln and Guba, 1985, p. 314).

The credibility and trustworthiness of the qualitative data are also strongly related to the interview analysis, as "we tend to see what we want to see, and hear what we want to hear" (Dey, 2005, p. 230), so that initial assumptions based on the analytical framework, for example, may lead to confirming the previous dispositions of the researcher (Miles and Huberman, 1994). Following Tracy (2010) and Creswell

and Miller (2000), the main strategies used in this research to mitigate these effects were to allow enough space for self-reflexivity between the three fieldwork waves, to allow for surprises going beyond the initial ideas and assumptions of the study, and to provide thick, transparent descriptions of the findings throughout all three results chapters. Plenty of time for data analysis during and after the qualitative fieldwork allowed for extracting meaning in the form of narratives, linking the findings to the theory and deriving implications (Rubin and Rubin, 2004).

It would be beneficial to measure actual activities and outcomes, for example in experiments. Both qualitative and quantitative data may be socially constructed (Bryman, 2008), because respondents are not so much "treasuries of information awaiting excavation" as they are "constructors of knowledge in association with interviewers" (Holstein and Gubrium, 2004, p. 141). Any human encounter beyond the research situation may be enacted performances comparable to theatre plays (Goffman, 1959). In the interview, the researcher, the participant, their relationship, and assumptions about their backgrounds may affect findings (DiCicco-Bloom and Crabtree, 2006; Ribbens, 1989; Herod, 1993). For example, the findings may be influenced by interviewer effects in relation to the interviewer's gender (Rubin and Rubin, 2004; Herod, 1993; Padfield and Procter, 1996), age and lifestage (Ribbens, 1989; Manderson et al., 2006), ethnicity (Manderson et al., 2006), social class (Manderson et al., 2006; Gray, 2009), previous knowledge of the topic (Arksey and Knight, 1999), and professional experience, for example with respect to previous medical education or in terms of being a doctoral student versus a professor (DiCicco-Bloom and Crabtree, 2006; Rubin and Rubin, 2004).

Consequently, surveys and interviews do not necessarily capture what individuals actually do, but rather what they *say* they do. These may diverge for Internet skills (Hargittai and Shafer, 2006; van Deursen and van Dijk, 2011a) and health perceptions as a sensitive topic (Krumpal, 2013), as explained in subsection 2.1.2. Then again, individuals' *perception* of their motivations, enablers and health outcomes are paramount, as these determine self-efficacy to achieve outcomes (Bandura, 1986, 2004), and shape subjective outcomes such as perceived health (Rosenstock et al., 1988), as argued in the literature review (see sections 2.1 and 2.2). Ultimately, this also relates to the underlying research paradigm: as it is relevant what individuals *say* and *think* they do, interviews and surveys assess their subjective perspectives while facilitating a certain generalisability, thereby creating the most relevant insights into Internet-based health outcomes in the spirit of the pragmatic paradigm (see section 3.1).

Triangulation of both approaches still remains challenging due to ontological and epistemological differences (Morse, 1991; Tashakkori and Teddlie, 1998). Whereas quantitative approaches encourage the finding of a single truth, qualitative approaches view the world as consisting of crystals "that reflect and refract, creating ever-changing images and pictures of reality" (Denzin and Lincoln, 1994, p. 482). This complicates merging data, interpreting divergent results, and determining the weight of each method (Mitchell, 1986; Patton, 1999). However, there are commonalities: both the quantitative and the qualitative data may be seen as representations of individuals' accounts, which enable the finding of common themes from individual voices. In that sense, moving between both data sources provided more holistic views of individuals' decisions and their thought processes: asking interviewees to rate their health on a scale from 1 to 5 led to a better understanding of the corresponding census item, while using the quantitative data to confirm the cluster allocation strengthened trust in the findings (see subsection 3.4.2). Consequently, as shown throughout this chapter, the mixed methods design not only helped to triangulate, complement and develop the research (Greene et al., 1989), but also enhanced its validity and reliability (Creswell and Clark, 2011; Tashakkori and Teddlie, 1998).

3.5 Ethical considerations

This research was granted ethical clearance by the Central University Research Ethics Council in March 2014, with two ethical issues being raised by this research: SM invariably leads to joining datasets without informed consent, and human participants' identities need to be protected (BSA, 2002; ESRC, 2015b).

For the first challenge, aggregation and anonymisation do not guarantee protection (Ohm, 2010), and combining datasets may pose threats to anonymity (Eynon et al., 2008). This is problematic as datasets such as the census and HES data were not collected for research purposes, and the public may have concerns about putting this data to new uses (Stevens and Laurie, 2014; Connelly et al., 2016). However, simulated microdata from aggregated records may be a more privacy-preserving way of generating data than collecting individual survey data (Hermes and Poulsen, 2012). In addition, HES data has only been released in the necessary aggregate form. Its appropriate use in terms of data privacy (not being able to identify individuals, and using the data for appropriate purposes) and data security (having the measures in place to prevent data leakage) was secured through the HSCIC's Data Access Approval Group (DAAG).

To protect participants' anonymity and ensure confidentiality, data analysis was based on the interview transcripts only (Bryman, 2008); these were separated from the contact data used in the recruiting process. As the combination of gender, age, education, socioeconomic status, long-term conditions and OA – with further details inferable from the quotes (DiCicco-Bloom and Crabtree, 2006) – might identify individuals, serial numbers contained in the publicly released OxIS dataset are omitted, and geographic location is reported only at the level of local authority districts (the smallest in this research contains 87,300 individuals). Further detail was only required for the Isle of Wight in subsection 6.3.3, which, however, does not specify exact locations either. As the type of locality and the OA classification are provided, this enhances participant protection without reducing the analytic value. Non-essential details, such as occupation or number of children, have been altered slightly to disguise participants' identity (Weiss, 1995). Vulnerable groups, such as children below the age of 18 or those with a lack of capacity, were not interviewed.

All interviewees received participant information sheets about the aim and scope of this research, and gave written consent to participating and being recorded (see Appendix XVI, p. 346), with the exception of the telephone interviews, where participants explicitly stated their consent orally. In addition, all interviewees were given the researcher's contact details, and could provide their email or postal address to be sent a short summary of the research findings (see Appendix XIX, p. 357), which are presented in full detail throughout the next three chapters.

Results 4

Health information for (non-)users

Kathleen is a 75-year-old retired civil servant who moved to King's Lynn & West Norfolk a few years ago and has never used the Internet. When meeting her for the interview, Kathleen took the participant information sheet, read it carefully and with interest, and then placed it on a small but substantial stack of printouts that she kept on a board below the table. It was not the first health-related sheet in the collection: one was on back problems, another one on the differences between physiotherapy and osteopathy, another one on the side effects of and alternatives to Warfarin, and a few more. Some of them were folded twice. Some were just printed as text, while others had the NHS logo on them.

Kathleen explained that she had obtained some of these from the local library, where she would regularly go to ask the librarian for information and have it printed to read. A few sheets were from a friend in London, whom she could phone and who would send her an envelope with printed information by post a few days later, though – as Kathleen later highlighted – her friend would go on the NHS website, not "on the Internet". None of the printouts were from her GP, but Kathleen valued that he also found out information about new treatments online upon request. Kathleen reasoned that she does not really need a computer, because she had various ways of obtaining information to learn more about treatments and make decisions.

The interview with Kathleen vividly illustrated how for health information, the boundaries between Internet users and non-users are becoming blurry. With health information being searched for by others and being easily passed on as printouts or orally, individuals can access it without necessarily using the technology them-

selves. In that sense, the access divide is inseparably linked with individuals' social environment (Warschauer, 2004; van Deursen and Helsper, 2015), as Kathleen's diverse network of individuals used for different purposes and on different occasions shows. At the same time, these blurring boundaries also go the other way: being a user does not necessarily imply being able to make meaningful use of the content obtained online (Warschauer, 2004), and individuals' use of the Internet may differ significantly between different kinds of users, as captured in the concept of the usage divide (DiMaggio and Hargittai, 2001; van Deursen and van Dijk, 2014).

With information as one of the cornerstones of the information revolution to which the NHS aspires (Department of Health, 2011), it is important to better understand how health information is obtained and used by different Internet users and non-users, as captured in RQ_1 and analysed in this chapter. While the known facets of digital inequalities help to develop a multifaceted understanding beyond the dichotomous use versus non-use, the qualitative data indicated that they tell only part of the story. Kathleen, for example, is officially a non-user, but based on her learning attitude, technology attitude, and view of the Internet as a useful information source, she may be different from other non-users – with implications for how the Internet affects her health service use and health perceptions. Drawing on the analytical framework and the qualitative data, this chapter argues that motivations and enablers may be seen as dimensions of different types of health information-seekers, and also shape the outcomes of Internet use. The resulting typology of health information-seekers provides a more appropriate way for describing individuals' orientations than the user/non-user split, and is a valuable frame for conceptualising Internet-based outcomes in relation to perceived health and health service use.

This chapter starts with a brief overview of these two health outcomes in section 4.1: their quantitative descriptive statistics and qualitative factors of relevance for their measurement throughout all three results chapters. Focusing on RQ_1 for the remainder of the chapter, section 4.2 analyses the two health outcomes in relation to all three user concepts derived from digital divide literature. Section 4.3 turns to the qualitative data, demonstrating how individuals' motivations and enablers are important for understanding how individuals use the Internet in a health-related context. Drawing jointly on the qualitative and the quantitative data, section 4.4 uses this as a starting point to establish a typology based on these motivations and enablers, which helps in the development of a more nuanced understanding of the outcomes of Internet use and online health information in section 4.5.

4.1 Overview of perceived health and health service use

Both health outcomes are spread across the range of values based on the quantitative data created through SM (see section 3.2). The range for perceived health is slightly smaller than the 1–5 scale since average ratings were taken to construct the variable (see subsection 3.2.2 and Appendix VI, p. 306). Both outcomes are skewed, as most individuals rate their health as (very) good, and do not go to hospital frequently.[10] Table 10 summarises the descriptive statistics for both health outcomes.

Table 10: Descriptive statistics for outcome variables

	Perceived health	**Health service use**
Range	1.87–4.80	0.00–12.00
Mean	4.12	0.28
Median	4.20	0.19
Standard deviation	0.32	0.44

The qualitative research confirmed that both perceived health and health service use are concepts that individuals can relate to and consider relevant. Health service use in particular was easily understood, both in terms of visits to the GP and to hospital, with the caveat that nearly all interviewees could not necessarily remember how many times they had made use of health services within the last 12 months. This demonstrates the usefulness of obtaining this value from the records.

For perceived health, most interviewees summed up their health problems and discounted their evaluation for their ailments:

> *Other than the fact that I'm getting a bit older, I'm extremely short sighted and got a cataract, other than that, I'm really quite well. And I have a little more weight than I should have [...] so I'm 4 out of 5, I suppose.*
>
> – **Anthony, 58**

However, individuals demonstrated a higher variation in terms of their understanding of the concept. Its subjective component was evident as individuals chose different yardsticks against which they judged their health, resulting in different views about what "perfect health" meant. Kathleen, for example, differentiated between her general and current health rating, factoring in her health problems:

[10] As discussed in subsection 3.4.3, the non-normality of residuals was fixed for all regressions throughout all three results chapters – see Appendix X (p. 328) for details.

> *I would say round about 3 to 4, generally. But it's about 2, I'd say, at the moment. It's not so good – it's not been so good this year.*
>
> **– Kathleen, 75**

Some older people generally accepted some health problems in their evaluation of their health, while some younger interviewees deducted points based on not doing enough sports and fitness-related activities (Blaxter, 1990):

> *I think for our age, we must be on 5s!*
>
> **– Karen, 69**

> *Hopefully I'm quite far up the scale: 4? I sometimes don't do as much as I could fitness-wise, otherwise I'd be a 5. But generally I've always been in good health.*
>
> **– Martha, 34**

Similarly, rating one's general health does not take into account the differences between physical and mental health (Johnson and Wolinsky, 1993). Women in particular rate their health partly based on their psychosocial wellbeing (Blaxter, 1990), but Ian also stressed this distinction when rating his health:

> *Physically I give myself a happy 5, I'm always blasting at the gym. Mentally, I'm a dreary 2. And compared to a couple of years ago, I would have said 0, so things are brightening up.*
>
> **– Ian, 42**

Ian's statement shows that comparisons to the past also play a role. Moreover, social comparisons between people influence how one would judge one's own health, as evidenced by Catherine, who was not able to walk without crutches, and Wanda:

> *About 4, I just never [let it] get me down, that's it. There's no point moaning about it. [...] The lady next door is in a worse state than I am.*
>
> **– Catherine, 77**

> *I've got a lot of friends who are [...] older than me, and a lot fitter. So that's what I'm comparing myself to, if you like. And they are far more active, they do Zumba and lots of other things. But I don't know that I could necessarily be doing that.*
>
> **– Wanda, 58**

What role the Internet plays in these social comparisons will be detailed in section 5.2. For now, the point is that how individuals rate their health adds variations to be taken into account throughout all analyses and discussion of them (Chapter 7).

4.2 Three user concepts and health

For the Internet's effects on perceived health and health service use, each user concept (access, support, and usage) tells a complementary, but more nuanced story.

4.2.1 Access divide

A first analysis of the quantitative dataset created through SM, comparing only use versus non-use, confirms that Internet use and perceived health are positively related ($\beta = 0.031$), while Internet users use health services less frequently ($\beta = -0.027$, neither shown in table), after controlling for the sociodemographic covariates that influence Internet use and health (see section 2.4). This also holds true when employing the access user concept following Dutton and Blank (2011). Table 11 shows that the relationships to both health outcomes are stronger for next-generation users than for first-generation users (perceived health: $\beta = 0.067$ and 0.034; health service use: $\beta = -0.037$ and -0.017). Ceasing to use the Internet is barely related to perceived health ($\beta = 0.004$), and associated with more health service use ($\beta = 0.015$). As an exemplar, Table 11 also includes two regressions with the outcome values appended to OxIS directly (see subsection 3.4.3): while the coefficients differ, the general relationships are similar. For virtually all following analyses, only SM results are reported; all other OxIS regressions are included in Appendix XI (p. 334).

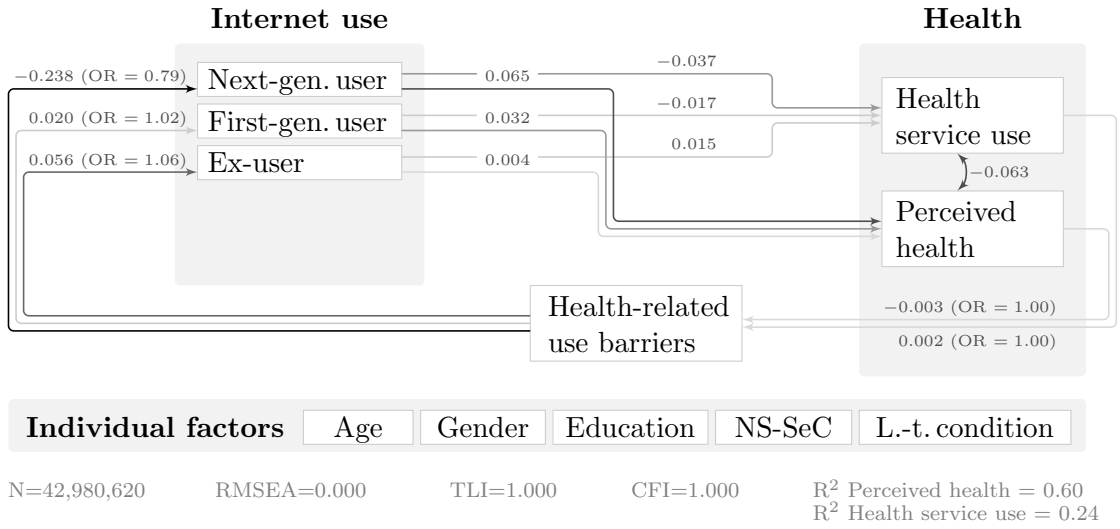
Table 11: Regression based on access divide

	Perceived health		**Health service use**	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
Next-generation user	0.067	0.061 **	−0.037	−0.035 *
First-generation user	0.034	0.020	−0.017	−0.038 *
Ex-user	0.004	0.023	0.015	−0.026
Never-user *(omitted)*	–	–	–	–
Age	−0.207	−0.159 ***	0.433	0.381 ***
Gender	0.094	0.126 ***	0.150	0.098 ***
Education	0.314	0.304 ***	−0.031	−0.016
NS-SeC	−0.083	−0.011	0.033	0.014
Long-term health condition	−0.462	−0.489 ***	0.030	0.032 *
Adjusted R^2	**0.60**	**0.55**	**0.24**	**0.18**
N	**42,989,620**	**2,150**	**42,989,620**	**2,150**
Largest CI	**6**	**3**	**6**	**3**

Note: Significance codes: *** < 0.001; ** < 0.01; * < 0.05. Throughout all analyses, HES is not adjusted for education, NS-SeC or long-term condition (see subsection 3.2.2; Appendix VI, p. 307), so that R^2s are lower for health service use.

Breaking down the relationship into a SEM (Figure 9) provides insights into potential causal links by incorporating the feedback loop with the concept "I have a health problem that limits my ability to use the Internet" (Q.D17 in OxIS, 2013, see Appendix IV, p. 298). While the relationship between Internet use and perceived health remains strongest for next-generation users (controlling for age, gender, education, NS-SeC and long-term conditions), those who agree that their health limits their Internet use are least likely to be part of this group ($\beta = -0.238$).

FIGURE 9: SEM based on access divide



NOTE: Some paths have been omitted to improve clarity: from the 5 individual factors *(age, gender, education, NS-SeC, long-term condition)* to all variables *(next-generation user, first-generation user, ex-user, health service use, perceived health and health-related use barriers)*, and covariances between each of the Internet use concepts *(next-generation user, first-generation user, ex-user)*.

The difference in first-generation users for perceived health may be due to their having their devices more readily available for health-related and other activities, with next-generation use being a "growing pattern of access [...] making the Internet more central to the life and work of networked individuals" (Dutton and Blank, 2013, p. 10). For example, Karen had plans to obtain a tablet and explains how this would allow her to use the Internet more frequently for health information:

> *Probably if you have something smaller, I would probably use it more. Like you would with a reading book. [...] By the time you've plugged [the computer] in, and it's all fired up, and I got to put passwords in, and I got to do this, and to do that... something smaller would be simpler.*
>
> **– Karen, 69, first-generation user**

Specifically for individuals whose health problems prevent them from using the Internet, having a mobile device may be beneficial. For example, Nathan, an avid Internet enthusiast, lives in a three-storey house and has trouble walking up the stairs to his computer. Using an additional mobile device helps him access the Internet despite his health problems:

> *I find [the iPad] much more convenient than the laptop or the desktop. [...] It's so convenient, you can just sit there, you don't have to climb the stairs, and you get the same information immediately. It saves me from walking, just pick it up, switch it on, and there it is.*
>
> **– Nathan, 90, next-generation user**

Moreover, disposable income is strongly linked to being a next-generation user (Dutton and Blank, 2011), which may explain part of the association between health-related use barriers and next-generation use. For Elizabeth, a lack of financial resources that came with her stroke prevents her from owning one (or multiple) devices:

> *I've only got my right hand working, my left side is paralysed, and my left leg is semi-paralysed, so in some senses [the stroke causes problems for Internet use], but it's finance mostly. [...] I mean you're not working yourself anymore, you get to live on benefits. You can't afford it anymore.*
>
> **– Elizabeth, 55, ex-user**

However, as in the example of Kathleen (introduced at the beginning of this chapter), Elizabeth has strong support from her family and the local library in terms of accessing the Internet. As a consequence, being an ex-user does not necessarily equate to having no Internet access, as will be explored in the next subsection.

4.2.2 Support divide

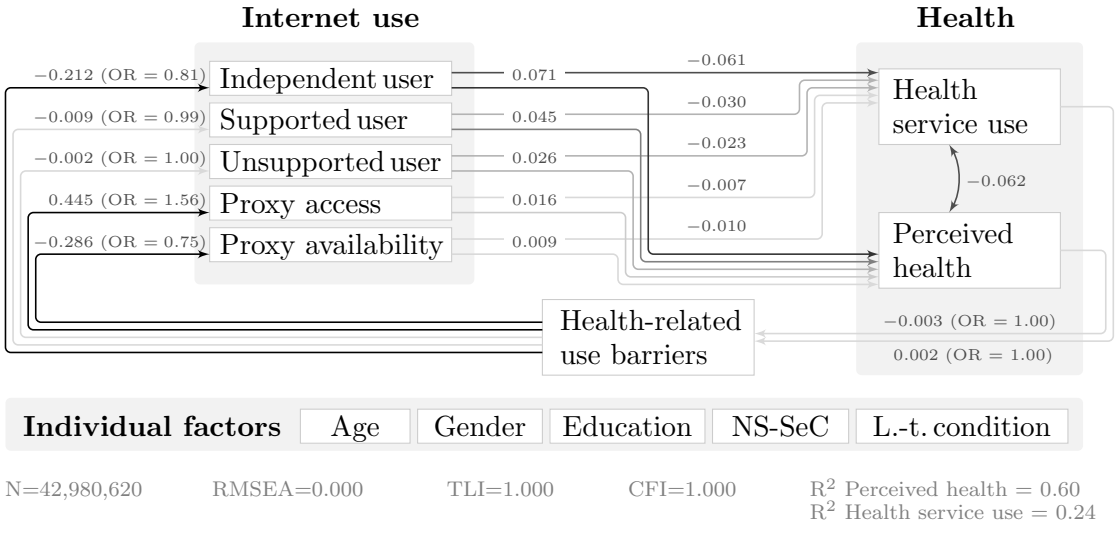
Conceptualising Internet use based on the support divide (Table 12) confirms the general differences between users and non-users, and leads to another interesting insight: for unsupported users, the association with both health outcomes is reduced close to the level of non-users ($\beta = 0.026$ and $\beta = -0.023$), supporting Warschauer's (2004) argument that those with unfulfilled support needs may not derive as much value from using the Internet. At the same time, perceived health and the reduction in health service use were highest for those who figured things out online without help ($\beta = 0.071$ and $\beta = -0.061$). For non-users, the relationship between proxy access and proxy availability is rather weak overall ($|\beta| \leq 0.016$).

Table 12: Regression based on support divide

	Perceived health	**Health service use**
Independent variables	β	β
Independent user	0.071	−0.061
Supported user	0.045	−0.030
Unsupported user	0.026	−0.023
Proxy access	0.016	−0.007
Proxy availability	0.009	−0.010
Fully excluded *(omitted)*	–	–
Age	−0.218	0.437
Gender	0.092	0.150
Education	0.316	−0.031
NS-SeC	−0.084	0.032
Long-term health condition	−0.462	0.030
Adjusted R^2	**0.60**	**0.24**
N	**42,989,620**	**42,980,620**
Largest CI	**8**	**8**

Figure 10 shows that having a long-term condition that reduces the ability to use the Internet was connected to lower chances of independent use ($\beta = -0.212$). Interestingly, proxy access was more common among those who said that their health represented a barrier to use ($\beta = 0.445$), while they were less likely to belong to the group who had someone available without making use of this proxy ($\beta = -0.286$).

Figure 10: SEM based on support divide



Note: Some paths, covariances and odds ratios are omitted (see Figure 9).

What united the supported and unsupported users were their generally low Internet-related skills. While the relationship between Internet use and health is stronger for supported users, Shirley explains that even with support, she would not use the Internet for health purposes due to her low skills, and also demonstrates low perseverance:

> *I ask a million questions and get help. I do [write it down], think that's easy, go to do it the next day, no, it comes up all different [...] so to learn to do big things on there, I'm just never going to get there. [...] But it's one thing where I quite happily think, no, I have to give that one a miss.*
>
> – **Shirley, 71, supported user**

For proxy users, the strong relationship between saying that health limits Internet use and proxy access may be explained by the link between wanting to use the Internet but not being able to (those who are not interested in the Internet, for example, may not say that health is the reason why they are not using it). Elizabeth, who stopped using the Internet following her stroke, has found several ways of accessing the Internet via a proxy because she has a desire to be online:

> *I'll ask someone, or I'll go to the library. It's not too bad. My son takes me in the car, so I can go in there. [...] You just have to find different ways. It's like a lot of things, you lose a skill, but [...] you adapt in so many different ways to accommodate for it. So I've used a lot of different ways to accommodate for my disabilities now.*
>
> – **Elizabeth, 55, proxy access**

Moreover, agreeing to the proposition that one's health condition limits one's ability to use the Internet may also depend on whether an individual is supported. Although Norman is a non-user with health-related problems, he does not think that his health limits his ability to use the Internet because, in his view, he *does* use the Internet since he delegates his requests to the most appropriate person:

> *If I want anything, I have somebody in Sheffield, he does my print-outs. If I read something, and I think I'd like to know about that, I phone him, and I say "I'd like you to find this for me", [...] he'll get the information and then send it down to me. [...] And [my friend's] wife is excellent on the Internet, she spends most of her time on the Internet. Anything on homoeopathy I would ask her. So different people for their different backgrounds, I suppose. [...] So, I do use the Internet!*
>
> – **Norman, 83, proxy access**

The results do not suggest that health issues were related to not having anyone available to help, either for users or non-users. Kevin, who suffered a stroke in

his 30s, was the only interviewee without any proxy available, as he was mainly in contact with (older) members of his stroke support group. While he stated that his health condition kept him from using the Internet on OxIS, the interview revealed that the primary reason for his non-use was his dislike of technology:

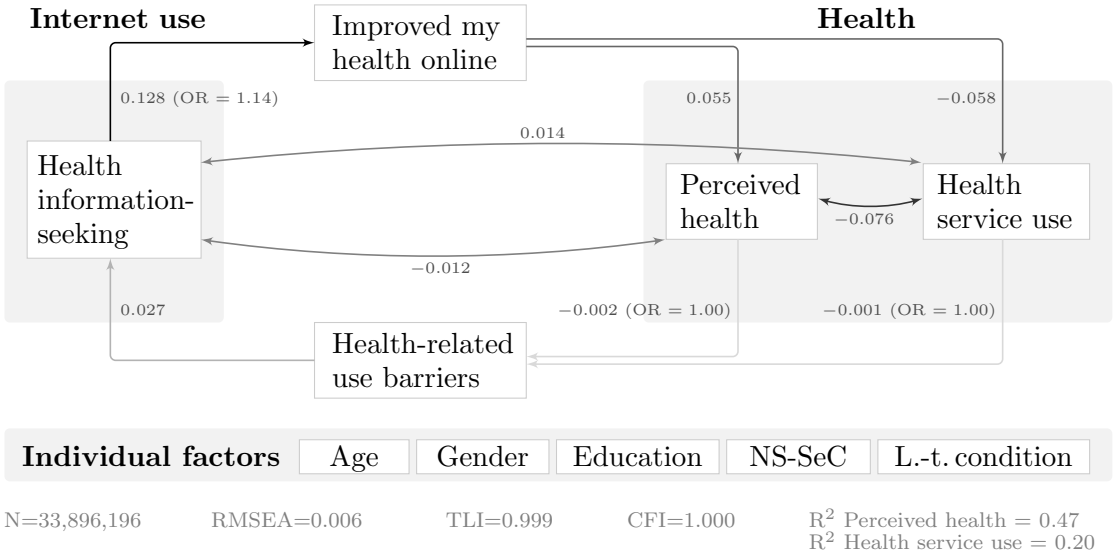
> *I don't use computers. I hate them. I'm on the phone, and I don't like phones. [...] Just don't like them. When we went to school, I didn't like computers. I don't like them. Full stop. I'm not bothered by them.*
>
> – **Kevin, 40, fully excluded**

The question of motivations and enablers for (health-related) Internet use salient in all four statements in this subsection will be taken up again in section 4.3.

4.2.3 Usage divide

Conceptualising Internet use based on the usage divide (DiMaggio and Hargittai, 2001; van Deursen and van Dijk, 2014) shows that health information-seeking was barely related to perceived health ($\beta = -0.005$) and health service use ($\beta = 0.009$) while controlling for demographic covariates (not shown in the table). However, incorporating the causally informed variable "Did you ever find anything online that helped improve your health?" (QC.22d in OxIS, 2013, see Appendix IV, p. 298) in Figure 11 shows that finding something health-related online was related to higher perceived health ($\beta = 0.055$) and lower health service use ($\beta = -0.058$).

Figure 11: SEM based on usage divide



Note: Some paths, covariances and odds ratios are omitted (see Figure 9).

The link between finding something online that improved one's health, higher levels of perceived health and lower levels of health service use is also reflected in the qualitative findings. For example, the Internet helps Miriam feel more positive about her state of health and puts her mind at ease, and at the same time reduces her health service use:

> *It puts my mind at ease. [...] For example the doctors said you got arthritis, and at first I was a bit... but when you read up on it, you think, well, it's not that bad. You can do this, you can do that. So it just puts your mind at ease, so you're not pestering the doctor all the time. [...] I wouldn't go [to the doctor's] as much through reading that. It's normal for these things to happen, just get on there, type it, [and] as long you don't need medication and stuff, there's no need to pester the doctor.*
>
> – **Miriam, 59, frequent health usage**

For the opposite direction, both health outcomes were barely associated with saying that health affected the ability to use the Internet ($\beta = -0.002$ and $\beta = -0.001$), which in turn was associated with higher levels of health information-seeking ($\beta = 0.027$). No interviewee said that his/her health conditions prevented him/her from seeking health information online specifically (in contrast to using the Internet more generally, as explained in subsection 4.2.1 and 4.2.2). Although it may seem counter-intuitive at first, those who experience health-related barriers to Internet use may also have more health-related questions to look up.

There is also a weak negative relationship between health information-seeking and perceived health that is not explained by the two causally informed variables: a lower level of health goes hand in hand with higher levels of health information-seeking ($\beta = -0.012$), which rather suggests that health causally affects Internet use. The qualitative data corroborated this assumption: many interviewees stated that specific health problems would trigger health information-seeking online, whereas there was little need to look up information if one was in good health:

> *No, [I didn't look up a lot], we didn't have the need to do that. There's also nothing really wrong with us, really, so were fortunate in that way.*
>
> – **Karen, 69, seldom health usage**

Then again, some relatively healthy individuals frequently looked up lifestyle-related information about cooking and exercising (such as Martha, Nicolas, Hugh and Isabelle). Interestingly, all of them rated their health status as 3 or 4 based on having a higher expectation of what makes up a 5 (Martha, p. 106), thereby defining their health status in relation to fitness (Blaxter, 1990). This alternative

explanation for why health information-seeking may be related to lower levels of perceived health will be followed up as part of the mechanisms between Internet use and health in Chapter 5.

For health service use, the same link holds true – an appointment may prompt looking something up for clarification after the appointment:

> *You don't always ask the questions at the doctor's that you should, so the information is there. [...] [It is] because you forget, it's as simple as that. It's "I should have asked about that" when you come out of the surgery.*
>
> **– Carol, 61, occasional health usage**

> *I search again, or look up what the doctor has given me or told me. [...] Because with hospitals, sometimes I can't understand what it is that they're saying. It's really odd when they're talking medical stuff.*
>
> **– Brenda, 40, occasional health usage**

On the other hand, Anthony – who generally says he does not use the Internet to look up health information – countered that he would always make sure that he understands everything explained to him after coming out of an appointment:

> *I try to make sure that it is pretty clear what they're saying to me. A lot of people [...] don't want to trouble the doctor. I've never been of that mind. If I'm going to see him or her, it's because I need their expertise, and so I don't want to come out more confused than when I went in.*
>
> **– Anthony, 58, seldom health usage**

The results presented throughout this section indicate that there is some relationship between Internet use and both perceived health and health service use. Part of the relationship may be explained by health affecting Internet use, as health problems may keep individuals from accessing the Internet (see subsection 4.2.1), increase the odds of being a proxy user (see subsection 4.2.2), and trigger health information-seeking online more generally (see subsection 4.2.3). However, as it relates to the main area of interest of this research, part of the relationship may be explained by Internet use affecting health, as indicated throughout the quantitative models and supported by the qualitative data in this section. The qualitative data also showed that the links by which Internet use may affect perceived health and health service use are complex and multifaceted, and apply differently for different individuals, as will be followed up in Chapter 5.

Although the user concepts derived from the access, support and usage divide are a useful step towards understanding the relationship between Internet and health, the qualitative data has shown that the reality is more complex. Over the course of

the interviews, the findings increasingly indicated that individuals could be grouped into different types in terms of their general and health-related Internet use, which in turn affected how the Internet shaped their health perception and health service use. Analysing the commonalities and differences between these types has revealed that the distinctive features for each type centre around antecedent factors in the analytical framework: motivations for general and health-related Internet use, and enablers for doing so (or a lack thereof). Both motivations and enablers will be explored in the next section.

4.3 Antecedent factors for health-related Internet use

The example of Kathleen, introduced at the beginning of this chapter, demonstrates the joint relationship between motivations and enablers. In terms of enablers, she had neither a computer nor the skills to use one. However, her motivations offset this lack of enablers: she thought that the NHS website was a fantastic resource for health information, and therefore found various ways to access it. This did not necessarily affect her health perception, but shaped her health service use, both in terms of discussions with her doctor and the treatments she had.

Section 4.4 and 4.5 will discuss these implications, both for Kathleen and five other interviewees introduced as representatives for the typology of health information-seekers, which serves as a lens on the relationship between Internet use and health. To establish this typology, this section provides insights derived from the qualitative data on both motivations (subsection 4.3.1) and enablers (subsection 4.3.2) for using the Internet for health-related purposes, and what quantitative constructs are available to reflect these in the secondary data (subsection 4.3.3). While the quantitative data does not do justice to the detailed qualitative findings, this enables the typology to be derived jointly from both data sources.

4.3.1 Motivations

Motivations for accessing health information reflected the Internet's affordances and the results of previous empirical studies (Powell et al., 2011; Ybarra and Suman, 2006; Ziebland et al., 2004; Chapple et al., 2012). From the interviews, five main motivations were identified: convenience and speed of access at all times; preparing for an appointment or treatment based on the breadth of information available; "translating" health professionals' advice through non-medical terminology online; building up further knowledge; and connecting with others to get peer advice.

These motivations do not necessarily reflect the intended and unintended benefits or drawbacks that individuals actually obtained (Brown and Walsh-Childers, 2009), but rather focus on *why* they used the Internet in a health context. Despite initially not asking directly for health information to avoid leading the interviewee (see interview guide in Appendix XVII, p. 349), all interviewees related "health and the Internet" to health information, and only later reported other ways in which the Internet affected perceived health and health service use (see sections 5.1 and 5.2).

The first area – **convenience and speed of access at all times** – was frequently mentioned as the main reason for health information-seeking, especially in terms of making a decision about whether to see a doctor. Hugh, for example, states that in virtually all cases, the Internet would be his first port of call:

> *I would probably always first enter the symptoms on the Internet, definitely before calling a doctor's appointment, for most things – to save time and curiosity.*
>
> **– Hugh, 45, user**

In particular with long waiting times, which the interviewees in the Birmingham and South Norfolk OA consistently mentioned, finding information online was seen as a quick way to assess whether there was a need to go to see a doctor.

> *I don't really go to the doctor's if I can help it. [...] I've always found it really useful [...] when you got the NHS self-analysing thing before you ring up for help. That's always useful, definitely.*
>
> **– Barbara, 41, user**

With respect to the time dimension, health information online was regarded as a source that would provide instant information. Even though Whitney did not see waiting times as a major hurdle, she stated that being able to find things out quickly helped her to avoid worrying:

> *[The main advantage is that] it is instant, because I worry about things in waiting. If I'm told I've got to have a procedure or something, I want to go – I don't want to go in six months time, or six weeks time, or in six days time, really I want to go instantly, but I know that's not possible.*
>
> **– Whitney, 60, user**

In addition to one's own time, the Internet may also help to avoid wasting the doctor's time. Several interviewees strongly expressed the view that being better informed helped them evaluate whether their problem was serious enough to justify seeking professional help:

> *It would help me to decide whether I actually need to go and see a doctor, or whether I think whatever it is can be treated otherwise. [...] I mean I'm not a person who likes to trouble doctors unless I absolutely have to anyway, because they are quite busy.*
>
> – **Wanda, 58, user**

A second range of health-related motivations revolved around **preparing for an appointment or treatment**. Several interviewees highlighted that preparation helped them to ask intelligent questions, appear more knowledgeable and get the most out of the limited time available in an appointment:

> *When my mum had the melanoma, I was on that computer the whole time looking up the specialists so we could ask questions, because you only have 20 minutes. And we came in completely well-prepared, to understand the information of what it could be, [and] what the future holds.*
>
> – **Helen, 43, user**

Several interviewees admitted they would not necessarily mention their online source, but rather use it tacitly to avoid giving the impression of questioning the doctor's skills, or to avoid risking dismissive comments from the doctor. Then again, particularly with respect to health services beyond primary care, Brenda's motivation for looking up information in advance was to be able to answer the receptionist's questions:

> *Sometimes it can be wise [to check] it on the Internet first. Because they do ask you these questions [...] "Have you seen someone first?", "Have you been to the GP?", "Have you looked this up?" They do ask, so I find this information first before I go into the walk-in centre or the hospital.*
>
> – **Brenda, 40, user**

A third related motivation was to **interpret professional advice** after appointments, as already mentioned in subsection 4.2.3 in the context of health service use triggering health information-seeking. This may relate to "translating" the information obtained from the health professional into simpler terms:

> *When you go to the hospital, they say "oh you got this, so and so thing", a C fracture actually. And as a layperson, you got no idea what that is, [so you] come home and google that quickly.*
>
> – **Matthew, 47, user**

The Internet also extends information through videos and images, which may help to clarify a condition or a certain procedure, such as an arthroscopy:

> *It shows you the picture on the Internet, it says actually where it is and [...] where they go, and it shows you all the little instruments used, there's been a cutter that was that big [...] it's just nice to look things up.*
>
> – **Miriam, 59, user**

Finally, some people also used the Internet to check if the doctor's decision was correct (Wilson et al., 2007). Barbara, for example, looked up the tablets the doctor had prescribed her and decided not to take them due to potential drug interactions or side effects for her depression:

> *Doctors have prescribed me Diazepan, and [...] we had a look [online] about it, and it does say basically it's not good for people with depression. So I haven't taken them. [...] It's all good information, sometimes better than going to the GP. Because they sometimes don't know themselves and they have to look into their big fat book. Honestly, we do a lot of searching on it [...] much better to google, I think.*
>
> – **Barbara, 41, user**

A fourth motivation for health-related Internet use included a desire to **build up further knowledge**. For example, with conditions such as MS, the Internet may have more up-to-date information than the GP can offer:

> *I looked up recent research for MS. I look up drugs that are available, if drugs are prescribed I look up side effects, to see what the side effects are. [...] The NHS tends to be lagging behind. It's often drug companies that [...] have done the research. The NHS gives very basic information, just [...] pure factual information.*
>
> – **Whitney, 60, user**

Several interviewees expressed an interest in learning about health in general, both with respect to long-term health conditions and more acute problems. Ahmad explains that he wanted to learn more in order to reduce future damage, while also emphasising his general desire to understand more:

> *I'm a curious person. Even though I've met the doctor, it doesn't mean that I can say "now I know". Even when I met the doctor, I always try to find more information about my problem – how to prevent this in the future. [...] I always want to know how to be more healthy, to be fit all the time, and I always look up on the Internet all the information.*
>
> – **Ahmad, 23, user**

Interestingly, this did not relate just to the medical conditions applicable to interviewees – some expressed a strong interest in learning about medical developments

more generally. Michael, for example, was fascinated by health topics and wanted to learn more simply out of curiosity:

> *I looked at [how to] stop smoking, and the impact that people can have on it. [...] And I was just looking up on how people see the world and if that actually affects your health if people see it differently. [...] I found that really interesting. Mental health, there's many things I looked up, I'm quite fascinated by that as well. [...] Things like that, I enjoy all that.*
>
> **– Michael, 22, user**

Finally, while the majority of interviewees talked about finding information online – and potentially stumbling across user-written information – a few individuals also mentioned **connecting with others to get peer advice** as a specific aim of going online. Danielle uses the Internet to simultaneously communicate with all her friends with children, rather than to contact someone specifically:

> *My youngest son had sickness and diarrhoea, so I was looking into that to see [...] when do I need to actually get him seen, how long is it before I need to start worrying, what are the symptoms, that sort of thing. And then I used Facebook actually and said to [all other] parents "Any advice?" [...] I have got lots of parents as friends. And then they've got obviously their friends who are parents as well, so it's just kind of like a mini mums.net on Facebook.*
>
> **– Danielle, 30, user**

Beyond the amplification of communications online, Whitney also specifically highlighted the value of user-written information and experiences for conditions such as MS:

> *You can see their experiences. Because if someone's actually tried it, and they sort of say what their side effects are, how it has affected them – they tend to know. Certainly with MS, it's one of these things that your GP doesn't know very much about. To be honest, they're not well informed: so many different aspects of MS, so many different forms, so many different symptoms! Basically the people that know most about it are the people who got it.*
>
> **– Whitney, 60, user**

Isabelle also highlighted the value of crowdsourcing opinions by reading user-written information, which would help in interpreting official information online, particularly when results from multiple sources match (Nettleton et al., 2005):

> *It's just to see their opinions, other people's opinions, if their opinions are the same as the other person's. [...] If you read something, and you*

> *read it two or three times, the same thing, you would be more inclined to believe that, than if you had only seen it once. [...] Even if it wasn't professional, you would tend to go with it more, if you see a lot of people writing.*
>
> **– Isabelle, 38, user**

These motivations show a breadth of reasons why individuals used the Internet in the context of health. However, various individuals had no motivation at all, and said they were not interested in using the Internet or had no usage opportunities (in general or for health purposes), as will be discussed in the next subsection.

4.3.2 Enablers

Some of the enablers for Internet use are interrelated, with low levels of interest often being related to not having a device and not having the skills, and vice versa (van Dijk and Hacker, 2003). This subsection disentangles their relationship for the specific context of health information-seeking.

In the context of health-related Internet use, **skills** in the sense of health literacy are an important enabler for the finding and evaluating of information. Several interviewees explained how they took account of the source in their evaluation of trustworthiness, and how they corroborated different sources to find answers:

> *If I get the odd person saying something really like over-the-top, I think I'm quite level-headed so I don't think "oh, that's what it means", I just think "oh, what an idiot" or something like that.*
>
> **– Martha, 34, user**

More generally, skills may relate to the technical skills of using a computer (van Deursen and van Dijk, 2011b). Despite having access, Shirley says that she would find both of these skills-related challenges difficult, and hence would rather not use the Internet for health:

> *We're on the Internet, but you couldn't say we use it. It uses me more than the other way around I think. [...] I'm not wise up enough to know how to do things, so [...] certainly don't do anything to do with money or health or anything else on it. So you could almost say: I'm not a user.*
>
> **– Shirley, 71, user**

In addition, those users without their own devices said they were worried about doing something wrong that might affect the other (more advanced) users of the machine. This view was expressed particularly by older women, who often rate their skills as lower (Hargittai and Shafer, 2006; van Deursen et al., 2011):

> *There's people that are brighter than I am, who can cope with that probably better [...] I mean, I am interested, but I'm always a little wary of it because [my husband] does such a lot of his things on it, and I keep thinking, I don't want to do anything, I don't want to wipe his stuff off.*
>
> **– Karen, 69, user**

Karen's statement shows the link to owning **devices** as an enabling factor. As she said earlier (p. 108), the kind of device matters as well, with mobile devices being more accessible. However, several interviewees also argued that for longer research on a topic, they would prefer to have a "proper" device:

> *If my granddaughter brings her laptop and I can use it, it's easier and bigger and I can see it better. So that's why I would prefer a laptop or proper computer [rather] than a smartphone.*
>
> **– Wanda, 58, user**

At the very extreme, someone may own no device to access the Internet. For example, Natalie needed to ask a family member to look up information for her because she had neither a computer nor a phone to access the Internet:

> *We just never got round to using it. It's expensive, we just couldn't afford them at the time. [...] But luckily we do have [my partner's] sister's to use if we do need it.*
>
> **– Natalie, 31, non-user**

Finally, in some areas, the kind of connection may also prevent individuals from looking up health information:

> *I'm just using it more here, because I got such a good signal here. [...] When we were still living in [a remote rural village], I wouldn't have thought of going on the Internet to look up things because it's such a quick and easy thing just to pick up your phone, type it in and it's there.*
>
> **– Isabelle, 38, user**

A third enabler relates to **interest** in using the Internet in general and for health information. For example, George generally was not interested in "high velocity stuff", despite using connected national databases before retirement as part of his job as a police officer:

> *I don't need it. I know what I need to know. If I want to find something out [...] I'm not bothered by looking at the high velocity stuff. That's what I call it, high velocity. These new things... I'm a dinosaur.*
>
> **– George, 69, non-user**

Individuals may also lack interest if there are others who can use the Internet for them (Reisdorf et al., 2012), which links back to the support divide discussed in subsection 4.2.2. This intersects with skills, but Whitney also highlights that her husband, who has MS, would not be interested in using the Internet:

> *There are people, obviously, whose family have made them do it. [My husband] doesn't, he's not interested, he has been on computer courses and all sorts of things – still doesn't bother. He isn't particularly interested because I can do it, it's easier for someone else to do it, but that's often the case.*
>
> **– Whitney, 60, user**

Specifically for health information-seeking, Sarah thought it was not "[her] job to look things up" (see discussion of the influence of socioeconomic status in section 6.1). She also exhibited a low interest in the Internet and her health more generally:

> *If I feel I need to go [to the doctor's] then I will, I usually wait a bit longer and if it gets worse then I go. It's like going to the dentist, you will wait, until it gets really, really bad you just don't go there. [...] No, I haven't yet googled it, if the problem continues then I might.*
>
> **– Sarah, 31, user**

Finally, the fourth enabler relates to **usage opportunities**. Although everyone will probably have health problems at some point, some people felt they did not need the Internet for health, for example due to a family member who provides them with information:

> *No, I've never looked up health information [...] My partner is a first aider anyway, so he knows what to do. And he's got a big book here, so if anything goes wrong we just look it up here.*
>
> **– Courtney, 29, user**

Some interviewees also found they did not need any health information as their doctor provided them with everything they needed to know. Catherine's statement also highlights the importance of language again, where the need to translate professional advice into plain English does not arise:

> *I'm lucky, I've got consultants that I believe tell you the truth. They don't fool you. I mean the doctor who'd done my back, he was absolutely fantastic. He told you in plain English what was wrong, and what he was going to do.*
>
> **– Catherine, 77, non-user**

Similarly, Nathan referred to his cardiac nurse as his main information source, and did not have any new health questions that he needed to look up:

> *I feel I know basically what's wrong with me, and what can be done about it. And that's probably because I know all those things, I'm not going to the Internet to find out. [...] I know where I'm going, and how quickly I'm going downhill. [...] So, I am content to know what I know through the nurses staff that I got or that come to see me, so I don't use the Internet for that purpose.*
>
> **– Nathan, 90, user**

To summarise subsections 4.3.1 and 4.3.2, looking at the different enablers showed that individuals may fall into user/non-user categories for very different reasons. Moreover, users may go online with different intentions in the context of health and more generally. Consequently, they may become non-users or proxy users for health information, and – as will be argued in section 4.4 and 4.5 – also experience different effects on health service use and perceived health.

4.3.3 Bridging qualitative and quantitative data

The quantitative data cannot entirely reflect these different motivations and enablers in all their facets. However, it is worth highlighting key themes that appeared in the presented motivations and enablers, which can be mapped to the available quantitative data. All resulting constructs are measured on 5-item Likert scales; an overview of these constructs and their corresponding survey items is provided in Appendix IV (p. 299).

First, several interviewees highlighted the **Internet's usefulness** in terms of convenience, speed of access and breadth of information available. The constructs around the Internet being an efficient means of finding information, making life easier and helping to save time (Q.I2a,d/Q.I8c in OxIS, 2013) reflect the level of convenience and the value of information that individuals see the Internet as providing. It would be desirable to include the idea of getting information from others (via crowdsourcing or by specifically looking for user-generated content) as well. However, while OxIS asks about the *uses* of user-generated content, there are no items that ask about the value of information from others, or using one's online network to obtain information. For this reason, the motivation of getting information from peers is not represented by a separate quantitative construct.

Reflected in the statements about interpreting and extending professional advice and building up further knowledge is individuals' **learning attitude** in terms

of understanding more or wanting to obtain more information to prevent future health problems, which is encapsulated in a construct around the desire to learn and understand new things (Q.I1e,g,i/SC.5h,i). The statements around building up further knowledge also reflect a certain level of **online enjoyment**: spending time online, reading everything about a topic, which is reflected in a construct around the Internet being a good medium to spend time on, and enjoying online activities (Q.I8g,h,j/Q.I2b). In addition, people's relationship with and attitudes towards medical professionals were also revealed in several statements. While it would be desirable to have more fine-grained quantitative data on the relationship with the GP, the quantitative concept of **trust in medical doctors** (Q.A7b) may be an acceptable reflection of individuals' views of the medical profession.

With respect to enablers, self-rated **Internet skills** (Q.C6) conceptualise the skills dimension. This relatively wide concept subsumes finding or evaluating information online, both of which are highly correlated with Internet skills ($\rho \geq 0.5$). Ideally, one would also include health literacy to account for the ability to find and carefully interpret medical information online; however, health literacy is not assessed in OxIS. For the attitudinal aspect of enablers around interest and usage opportunities, **Internet interest** (Q.A9) captures individuals' desire to go online. With low trust in technology being reflected in several statements in the enabler subsection (for example, the technology might break, or "high velocity stuff"), **technology attitude** (Q.B1a,c,g/Q.I1a-c), expressing general views of technology, is a useful construct to reflect enablers of general and health-related Internet use.

The last construct of **self-efficacy** also reflects the health behaviour literature (Bandura, 2004). It would be useful to have a construct that specifically relates to individuals' sense of being involved and thinking that they can actively do something to contribute to their health. However, the extent to which one can shape the direction in one's life (SC.5f) captures the idea of actively being able to do something about one's health, which surfaced in several of the interviews.

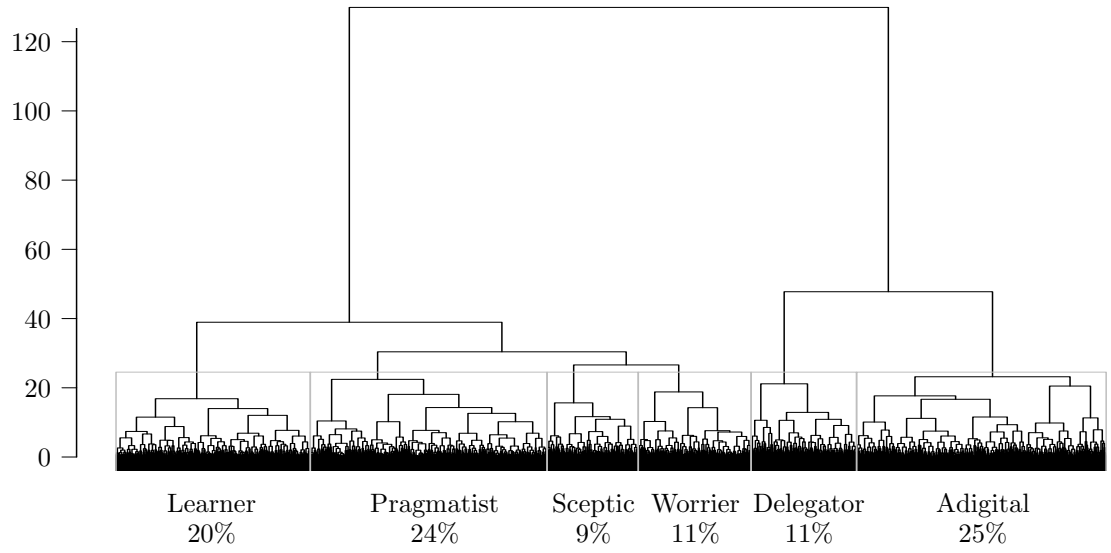
In general, it would be useful to include these constructs more specifically for health. However, based on previous research, health information-seeking was relatively similar to other informational activities online (Blank and Grošelj, 2014). Therefore it may be considered reasonable to use these general attitudinal variables derived from the qualitative data. The residual difference between qualitative findings and their quantitative representation that arises from using these constructs in developing the typology will be pointed out, where applicable, in the results presentation in section 4.4.

4.4 A typology of health information-seekers

Based on the qualitative notion that there are certain types of health information-seekers, which in turn shapes how the Internet influences their health perception and health service use, this section utilises the constructs derived from the qualitative data in subsection 4.3.3 to establish a typology of health information-seekers. The typology helps to delineate individual differences in terms of how the Internet affects the two health outcomes in this research, and thereby makes a major contribution to answering RQ_1. Rebuilding it with the quantitative data gives the typology a more solid foundation, as it makes it possible to evaluate whether the quantitative characteristics of the clustering dimensions are reflected in the qualitative observations, and allows for cross-checking whether the 31 interviewees recruited from OxIS fall into the same clusters based on both data sources. Moreover, by revisiting some of the analyses from section 4.2 separated by each type in section 4.5, this also serves to provide a quantitative answer to RQ_1, in addition to the qualitative findings.

The resulting clusters partially delineate the user/non-user split, although no clustering dimensions were use-related: the dendrogram in Figure 12 shows two large groups, with mostly users in the left part and mostly non-users in the right. Based on the dendrogram, different numbers of clusters would have been justifiable, but six were chosen based on the qualitative interviews, keeping cluster interpretability as paramount. For the final six-cluster solution, the average within-cluster distance

FIGURE 12: Dendrogram for hierarchical clustering of typology

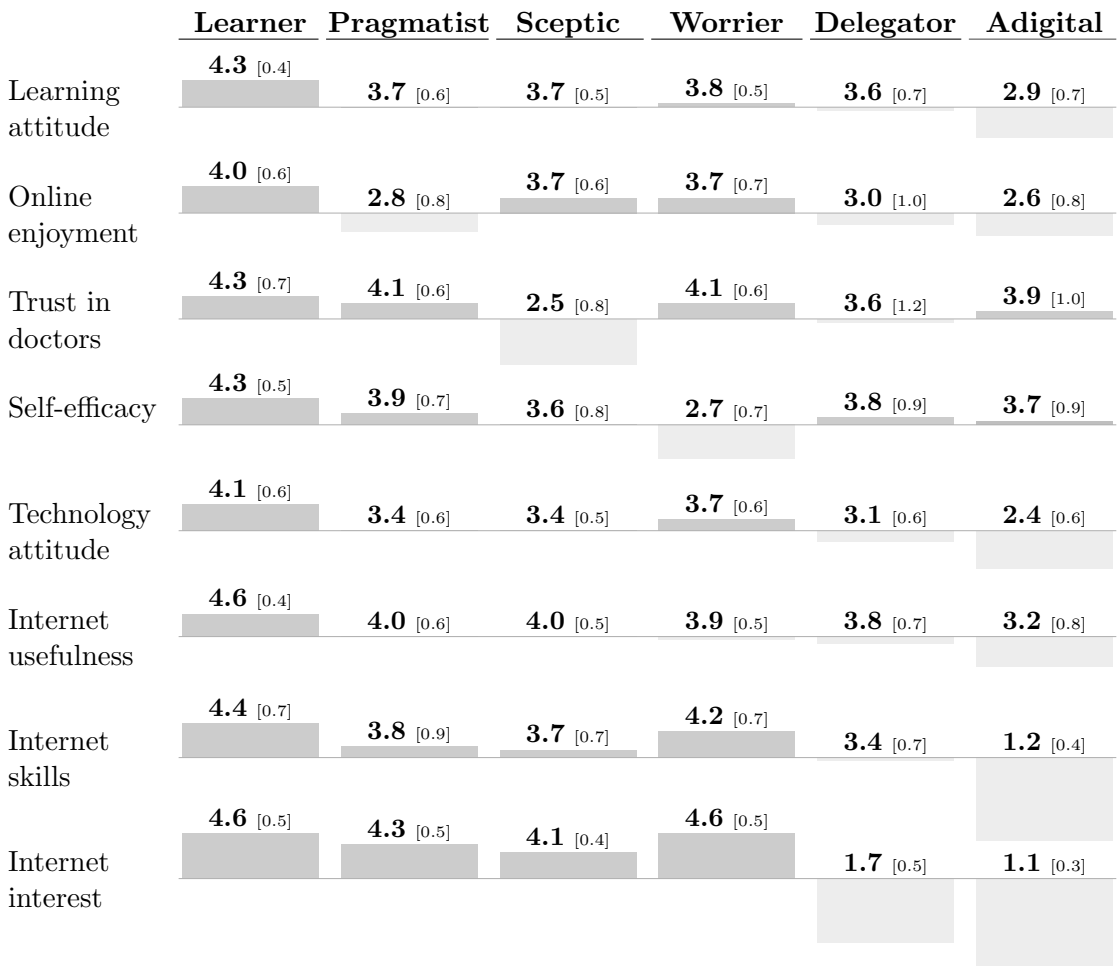


NOTE: Percentages are taken from weighted OxIS data.

is smaller (2.5) than the between-cluster distance (4.3). With an entropy of 1.7, the quality of the clustering solution corresponds to other models (Zhao and Karypis, 2004). The dendrogram also reflects diversity within each cluster: as one would expect from real-world social data, these types are not completely homogeneous.

The six resulting types were named the *learners*, the *pragmatists*, the *sceptics*, the *worriers*, the *delegators* and the *adigitals*. On the one hand, these names stem from the qualitative findings, which will be presented by introducing a representative of each of the six types subsections 4.4.2–4.4.7. On the other hand, their names reflect some of the salient quantitative characteristics of the clustering dimensions shown in Figure 13, which will be explained in more detail throughout the next subsection.

FIGURE 13: Typology of health information seekers – cluster dimensions



NOTE: All values follow the notation: mean [standard deviation]. The diagram shows the divergence from the mean. As all constructs are measured on 5-item Likert scales, they count with equal weight in the clustering (Gan et al., 2007).

4.4.1 An overview of the six types

In short, the **learners** primarily used the Internet to find out everything about their or others' conditions and enjoyed doing so, both before and after seeing a doctor. This is corroborated by the high values across all clustering dimensions. In contrast, the **pragmatists** tended to value the instrumental efficiency of the Internet, but rather used online resources as a decision tool to find out if they needed to see a specialist, but with low enjoyment experienced in finding out more. They often saw the Internet as a double-edged sword, as supported by their average values for technology attitude and Internet usefulness. While medical **sceptics** are similar to the pragmatists in many dimensions, they enjoyed trawling through online resources, but have a sceptical attitude towards doctors. Consequently, they primarily saw health information online as a way of ensuring that health professionals were doing their job correctly, and avoiding health service use wherever possible. Similarly salient for the **worriers** was the low level of self-efficacy, which links to interview insights about rather not wanting to know, and judging the possibilities of shaping their own health as faint. While the worriers are similar to the learners in many ways, several admitted that their inquiries often led to increased concerns. **Delegators** saw the value of the Internet and also health information online, but were restricted due to a lack of one or several enablers (see subsection 4.3.2). They would rather not use the Internet as a first port of call due to the efforts connected with doing so, and especially valued the comfort of delegating information-seeking online to others. Finally, the **adigitals** did not think the Internet was useful in general, and specifically in the context of health had strong views against using health information online; this was often combined with low interest and skills in using the Internet.

As in the real world, these groups are not mutually exclusive. For 6 of the 31 OxIS interviewees, the quantitative and qualitative classification did not correspond. These interviewees were manually reclassified based on the qualitative findings in accordance with the analytical strategy (see subsection 3.4.2), adopting the method by which those that originated from the same OA (without having been part of OxIS) were classified. The divergence can be traced back to various factors: changed attitudes towards the Internet between the OxIS survey and the interview, as reflected in a recent decision to quit all social networks (Sarah, reclassified from sceptic to adigital); recently changed attitudes towards doctors and the health system based on previous experience (Elizabeth, reclassified from pragmatist to sceptic); attitudes towards health information that differed from other kinds of information (Anthony, reclassified from learner to worrier; Norman, reclassified

from adigital to delegator); and a general mismatch between qualitative and the quantitative classification (Susan, reclassified from pragmatist to adigital; Wanda, reclassified from worrier to pragmatist – she also wanted to correct some of her survey responses). The implications of these mismatches for the interpretation of results and the validity of the typology will be discussed in subsection 7.5.1.

However, the overall fit between the qualitative and quantitative results confirms that the typology represents a useful systematisation of orientations towards health information-seeking, and shows how individuals perceive their health and make use of health services, as will be argued in section 4.5. Including only motivations and enablers as clustering dimensions does not imply that other variables – such as individual sociodemographic factors, long-term health conditions and other contextual influences – do not play a role. In fact, as Chapter 6 will elaborate, some of the motivations and enablers are linked particularly to age, education, NS-SeC and long-term health conditions. The reason for leaving them out as clustering dimensions was to derive the clusters as parsimoniously as possible, and rather evaluate the sociodemographic differences across types based on the resulting clusters.

Table 13 shows selected sociodemographic and use-related aspects that were not part of the cluster analysis. Interestingly, despite the user/non-user split between the first four and the last two groups visible in the dendrogram (Figure 12), these types cut across different user concepts. This implies that some users are closer to non-users in their motivations and enablers, and in their use and the effects of the Internet. Then again, some non-users belong to the learners, pragmatists, sceptics and worriers; this is particularly true for those who gave the lack of a device or skills as their main reason for not using the Internet, whereas the majority of non-users in the delegator and adigital groups cited a lack of interest as the main reason for non-use. The types also cut across those with or without long-term health conditions, as has also been reflected in the qualitative interviews (see subsections 4.4.2–4.4.7).

Moreover, all types include individuals who do not seek health information online, albeit to a lesser extent in the first four groups, which comprise more users. Based on the interviews, not using the Internet for health information-seeking was related to four factors largely observable across all six types. The first was having alternative informal sources (the interviewee's own or a family member's medical knowledge), such as Courtney, whose partner was a paramedic, or David, whose wife worked in a pharmacy. The second was relatively easy access to doctors, which manifested itself in convenient access (Nathan, p. 123), low waiting times for appointments, and good explanations from the doctor (Catherine, p. 122). A third revolved around

Table 13: Typology of health information seekers – other characteristics

	Learner	Pragmatist	Sceptic	Worrier	Delegator	Adigital
Demographics						
Age, $\mu\,[\sigma]$	37.8 [15]	46.2 [16]	43.1 [16]	41.9 [16]	46.7 [19]	66.5 [16]
Education, $\mu\,[\sigma]$	2.6 [1.3]	2.4 [1.4]	2.2 [1.4]	2.3 [1.5]	1.8 [1.1]	0.6 [1.1]
NS-SeC, $\mu\,[\sigma]$	2.2 [1.0]	2.0 [1.1]	2.3 [1.1]	2.3 [1.0]	2.4 [1.1]	2.7 [0.8]
Gender, % female	49	50	47	48	58	59
Long-term condition, % of all individuals	13	13	10	17	20	48
Access divide (Proportion of total group in %)						
Next-gen. user	83	61	62	66	34	1
First-gen. user	16	36	33	31	38	7
Ex-user	1	2	2	1	10	7
Never-user	0	2	2	1	18	84
Support divide (Proportion of total group in %)						
Independent user	75	60	48	42	44	4
Supported user	6	15	19	29	13	3
Unsupported user	19	21	21	25	15	1
Proxy access	0	2	3	0	8	16
Proxy availability	0	1	3	4	13	42
Fully excluded	0	1	5	0	7	35
Usage divide: Health information						
Seeking, Frequency	1.7	1.6	1.7	1.8	1.5	1.0
Seeking, %	71	70	56	79	54	32
Finding, %	52	40	43	28	28	11
Users: Online activities						
Social, %	93	83	85	86	51	3
Information, %	99	93	92	96	70	6
Entertainment, %	92	77	76	82	40	2
Transaction, %	99	96	95	97	71	8
Non-users: Most important reason for non-use (Proportion of total group in %)						
Lack of skills	0	41	0	23	0	8
No devices	100	57	55	62	29	10
Lack of interest	0	1	45	15	71	80
Usage opportunity	0	0	0	0	0	2
Health outcomes						
Perceived health, $\mu\,[\sigma]$	4.23 [0.3]	4.18 [0.3]	4.18 [0.3]	4.17 [0.3]	4.11 [0.3]	3.82 [0.4]
Health service use, $\mu\,[\sigma]$	0.19 [0.3]	0.25 [0.4]	0.23 [0.3]	0.21 [0.2]	0.27 [0.4]	0.50 [0.7]

Note: The outcome variables for perceived health and health service use largely reflect the sociodemographic group characteristics and should not be interpreted as one group being "healthier" or using health services less frequently.

low need due to no major health concerns (Karen, p. 113). Finally, the worriers displayed a strong normative sense that they *should* not use the Internet for health information (see subsection 4.4.5). However, the worriers also have the highest health information-seeking frequency across all groups ($\mu = 1.8$, Table 13). This may indicate that their online activities fuel their concern, which in turn leads to more health information-seeking (Fergus, 2013) – or relate back to interview data being socially constructed reflections of what individuals say they do (see subsection 3.4.4).

As an alternative way of delineating the types using the cluster dimensions (Figure 13) and further characteristics (Table 13), Table 14 presents a logistic regression on each type. It shows that some clustering dimensions predict belonging to a certain type, while sociodemographic features are less informative. It also illustrates that individuals are not deterministically assigned a type, but rather have a higher probability of falling into one type or another based on their antecedent factors. As the types were formed based on OxIS data alone, the analysis below only uses sample data, and therefore includes statistical significance measures to evaluate the confidence that the ORs are representative of the population.

Table 14: Logistic regression for types of health information seekers

	Learner	**Pragmatist**	**Sceptic**	**Worrier**	**Delegator**	**Adigital**
Independ. variables	*OR*	*OR*	*OR*	*OR*	*OR*	*OR*
Learning attitude	1.83***	0.94	0.55***	0.51***	1.44*	0.15***
Online enjoyment	5.32***	0.45***	1.07	1.13	1.35*	0.13***
Trust in doctors	2.76***	1.21**	0.10***	3.80***	0.97	0.72
Self-efficacy	5.14***	3.02***	0.44***	0.05***	0.81	1.14
Technology attitude	1.86***	0.54***	0.56**	0.79	0.82	0.58
Internet usefulness	1.80***	0.71***	0.43***	2.70***	2.68***	0.30**
Internet skills	3.99***	0.60***	3.04***	0.55*	73.69***	0.00***
Internet interest	4.14***	17.98***	14.15***	20.03***	0.00***	0.00***
Age	1.04	1.15	1.18	0.88	0.59**	0.82
Gender	1.06	1.00	1.06	0.90	1.25	0.94
Education	0.88	1.17	1.09	0.94	0.98	0.86
NS-SeC	1.04	0.98	1.10	0.98	0.83	0.78
Long-term condition	0.92	0.90	0.92	0.93	0.78	0.95
Pseudo-R^2	**0.73**	**0.48**	**0.62**	**0.71**	**0.77**	**0.98**
N						**2,150**
Largest CI						**4**

Note: Significance codes: *** < 0.001; ** < 0.01; * < 0.05; the Pseudo-R^2 follows Cragg and Uhler's definition.

The following subsections introduce each of the six types of health information-seekers, outlining their approach to health information online, how the Internet is likely to affect how individuals perceive their health and their health information-seeking, and some other salient characteristics for each of the types. As outlined in the analytical strategy, this section presents only one typical representative with the main characteristics identified for this type, as the key in any typology is that certain features occur within one single person rather than multivocality (Tracy, 2010). The final section (4.5) of this chapter features some more voices for each of the types, while Chapter 5 also includes further aspects of the six portrayed individuals that were not representative of the type as a whole.

4.4.2 The learners

The learners have a strong desire to understand health better. While they do use the Internet to make decisions about whether they need to see a professional, more than anything they use the Internet to learn about their health: before seeing the doctor, after seeing the doctor, and between appointments. Notably, this could be for their own health conditions, but also for the ailments of others in their social circle or neighbourhood (as will be covered in more detail in Chapter 6).

Miriam from Middlesbrough is an example of this group. At the age of 58, she is older than average ($\mu = 37.8$), near the average education (level 2, $\mu = 2.6$) and near the average socioeconomic status (NS-SeC 2, $\mu = 2.2$, Table 13). Having had a minor stroke and arthritis, she has long-term conditions, as does Ian, although this group also comprises those without chronic conditions, such as Michael or Ahmad. The primary value of the Internet for Miriam is that it allows her to better understand her conditions. After seeing the doctor, she would usually google the information she had picked up and read everything she could find about it, in line with the generally high learning attitude of the group of learners ($\mu = 4.3$, Figure 13):

> *They explained a lot more [online]. The doctors give you the basic information that you need to know, this is what it's called, [but] it explains a lot more on the Internet as to what the causes are, how you can reduce it and so on. [...] This is just when you don't get a full answer from the doctor about what this is, I like it to be expanded a bit more – whether it's good news or bad news.*
>
> **– Miriam, 59, user**

This is the case because doctors' time is limited, and having additional information makes her feel "on top of things":

> *I know the doctors have to see somebody in like 20 minutes, so they have a limited time for how much they can tell you. [...] I quite like looking at it – sounds a bit nosy, doesn't it? But I really like to look on it and just, and it keeps me on top of things, I suppose. And it makes me aware of what I need to look out for, that's good.*
>
> – **Miriam, 59, user**

Often, she would look up something before seeing a doctor in order to see if she needed to go, and afterwards again, indicating a strong sense to be involved in her treatment, for example through finding additional ways to ease her arthritis:

> *Sometimes I've looked before, if I'm not sure what something is, I'll type my symptoms in and I'll have a look. Because sometimes you think, do I need to go to the doctor's and I'll look it up and see what it is [...] Then usually I'll look it up again after I've been to the doctor's, just to see what you can do to ease it a bit, or if there's anything that can help you.*
>
> – **Miriam, 59, user**

Specifically in the context of her minor stroke, Miriam found that information online was easier to understand due to the use of plain language, and could be read whenever necessary. This helped her to internalise the information, act on it, and be relieved when she felt she had understood the doctor's advice properly:

> *Sometimes when you go to the doctors you get scared because you don't know what they're telling you, and then [online] it is in simple English, and [...] it just puts your mind at ease. [...] I've read it over and over again, so it has sunk in and it didn't upset me, and I think I know what to do, but when the doctors say it you think wow.*
>
> – **Miriam, 59, user**

Beyond the need to look something up, Miriam also shows enjoyment in being online and reading about medical conditions. The group of learners generally shows a higher trust in other people online than all other groups ($\mu = 3.2$) and thinks that the Internet is a good place to meet people ($\mu = 3.2$; both not shown in table). Beyond her own condition, Miriam is also very interested in other people's stories:

> *I do like to read up on these, so I see if there's something I haven't got – just see how people [...] deal with a situation and what the result is, because you can get six people with the same ailment, and they've all tried it differently. [...] I just find it interesting, I like to know what these people took, and sometimes I would go on for that, just to read about this, and it's just amazing if you see the different outcomes just by people taking what they have at hand [...] So I'm just nosy.*
>
> – **Miriam, 59, user**

The learning aspect goes beyond her specific condition, with Miriam also setting herself up as a "lay expert" for other people. This complements points from other learners such as Michael or Danielle, who highlighted how they could give advice based on information obtained online:

> *If somebody's brought something up that something's happened to them, I tell them that I would give them my point of view once I've read up on it on the Internet. I'll have a look on the site, just type the word in to reassure them a bit.*
>
> **– Miriam, 59, user**

Finally, more than specific advice, many learners were also particularly interested in building up knowledge, reading other people's stories and disease journeys, or finding advice on fitness and lifestyle. Miriam says that reading about other people's stories would recalibrate her perception of her own health:

> *When you read what other people are going through, and you're going through something similar, but they're at a different time, and now they're happy, because they've done this, I think I'm not on my own. I'm not the only one who's going through this, and I can come out on the other end. [...] And through looking at that as well, that's what makes me think, well you're lucky really [...] So it just made me feel like "stop whining about it and live your life". These people are a lot worse off – it just put things into perspective.*
>
> **– Miriam, 59, user**

In summary, many of the learners, such as Miriam, Ahmad, Michael and Danielle, emphasised that information online puts their mind at ease and makes them feel better about their health, thereby leading to positive effects on how they would perceive their health, while – as with the pragmatists – making, in their own view, their health service use more appropriate.

4.4.3 The pragmatists

The group of pragmatists primarily uses the Internet as a decision tool in order to decide whether it is worth seeing a doctor, both in general and during out-of-office hours. The group generally has a high appreciation of the medical profession, trusts doctors and official websites more than opinions voiced online, and values highly the opportunity to use online services, for example to request repeat prescriptions and make appointments, and potentially even have online consultations.

Hugh is the head of an international design company based in London, and lives in the borough of Hammersmith and Fulham. Aged 45, he is about average

in this middle-aged group ($\mu = 46.2$), with a high education (level 4) and a high socioeconomic status (NS-SeC 1) reflecting the relatively high levels of both for pragmatists ($\mu = 2.4$ and 2.0 respectively, Table 13). His primary motivation for looking up health information is to make a quick decision about whether to see a doctor or not, in line with the low level of online enjoyment of this group ($\mu = 2.8$, Figure 13):

> *I would use it just to kind of get a handle on whether it's kind of worth going to the doctor or not, not for kind of detailed self-diagnosis. [...] So I was to look up a complaint, and I was to satisfy myself that it was nothing to worry about for someone of my age, or my child's age... I mean, I'm not one of these people that likes to go to the doctor anyway, because it's the loss of working time.*
>
> – **Hugh, 45, user**

Depending on the specific circumstances, this may result in lower health service use, though in practice he often nevertheless has to see a doctor, especially if a prescription is required. However, in these cases, the Internet helped him to evaluate the urgency of the situation:

> *One of my children had impetigo. So when I saw the symptoms, I went online and [...] it was 90% certain that it was impetigo in my mind. So then I made a doctor's appointment, and she confirmed it and gave me antibiotics. I would have probably gone to the doctor anyway, because [...] although it's not very serious, it gets quite nasty. [...] But I guess it gave me a bit of comfort that this is all it was.*
>
> – **Hugh, 45, user**

However, Hugh would usually not take this information to the doctor. In general, the group of pragmatists has a high trust in doctors ($\mu = 4.1$, Figure 13), but Hugh also shows a high appreciation of professional experience based on his own job:

> *I might challenge [the doctor], but I let him to give his diagnosis first. I wouldn't go in there with it. I mean obviously I wouldn't expect someone to come in here and tell me, a client to tell me the correct way to detail. He'd certainly give me the brief, but not how to light a space, or to build the space. You have to rely on the expertise of the professional first, and then if you don't like what they do, then you [say something].*
>
> – **Hugh, 45, user**

Like many other pragmatists, Hugh does not want to share and discuss health problems via the Internet, and does not value non-professional advice online:

> *I think that's more to do with being British than the Internet, [but] I wouldn't share this. You might do if there was an expert, that's a different question. [...] On social media, I'm assuming you would put your problems out there like you would put a dump in your house out there, and see who's got an opinion on it. I wouldn't, I mean that's an extension of self-diagnosis, that's getting other neurotics to share your problem with.*
>
> **– Hugh, 45, user**

The group of pragmatists generally includes relatively Internet-savvy individuals in terms of general Internet skills ($\mu = 3.8$, Figure 13). Hugh explains that he thinks that most people today know how to navigate the Internet:

> *I think most people – and that's not just medical, that's just modern Internet usage – you have a street wisdom, don't you? You just know how to search through google and cut out the more commercial sites. You have a sixth sense for it.*
>
> **– Hugh, 45, user**

Like the learners, the pragmatists are also open to health services online, and value convenience of services. This is reflected in the relatively high proportion of users of information (93% of users) and transactional services online (96% of users, Figure 13). Hugh would also embrace medical services delivered online:

> *[I wouldn't have concerns], not if I knew the site was safe, that I was dealing with a professional and therefore by implication the people who set up that site had appropriate levels of Internet security [...] I wouldn't be any more concerned about privacy or security than with Internet banking.*
>
> **– Hugh, 45, user**

Like Hugh, many interviewees in the group of pragmatists showed a high understanding of the need for new approaches to care and health provision:

> *I'm all for putting less pressure on the health service through faster [and] more efficient forms of medical support for which you don't need face-to-face, definitely. [...] And therefore you should turn the wheel, there's nothing scary about turning the wheel. But as you know, in Britain people are very proud of their health service... [it's] sacred ground.*
>
> **– Hugh, 45, user**

It is worth noting that Hugh does not have any long-term health conditions, though other individuals who fall into this group – such as Wanda or Matthew – do. Despite these long-term conditions, however, both said that they would not be *"bothered with it so much"* or would not like to get too involved in their conditions.

4.4.4 The sceptics

In contrast, the medical sceptics see the Internet as a tool to solve health problems that doctors cannot deal with. Put differently, they are sceptical of the health system and exhibit low trust in doctors ($\mu = 2.5$, Figure 13) – but they embrace the Internet for health information. The medical sceptics mostly said that the Internet would lower the frequency of health service use for them, although they would still have to go there for prescriptions and procedures, for example.

Brian is an example of the medical sceptics. At 54, he is slightly older than the average of the group ($\mu = 43.1$), and represents the slightly higher share of males in this group (47% females, Table 13). Brian is from Birmingham and used to work in the middle management of a courier business. He does not think of doctors as a very useful resource, partly because getting appointments is difficult:

> *The doctors are more and more useless now as time goes on. It's very difficult to get appointments [...], quite often two weeks is the minimum, three weeks or more until you get the appointment.*
>
> **– Brian, 54, user**

On several occasions when Brian went to the doctor, he felt that the information that he had obtained online in fact alleviated his problem. In his view, it is essential that people look up information online to ensure a proper diagnosis:

> *I've been to the doctor [about my severe back pain], three or four times and he was suggesting, well try this, and try that, and they sent me for kidney tests, and that involved quite a few blood tests [...] but I was just reading about back pain on one of those forum things, and [...] someone was talking about the benefits of drinking more water. [...] And that was the answer! [...] So really, today, because it's so hard to see the doctor anyway, you're better off if you can [look it up online]. Years ago they said you shouldn't self-diagnose [...], but I think today, it's kind of essential that you look up many things yourself and then see.*
>
> **– Brian, 54, user**

The group of medical sceptics exhibits low trust in doctors ($\mu = 2.5$, Figure 13), which is reflected in Brian not believing his doctor, and rather measuring his blood pressure himself to show the doctor that there was something wrong. His last sentence also points to differences between "doctors in general" and "my doctor" (Hall et al., 2001):

> *I have high blood pressure, and I went to the doctor, and he said that's nothing to worry about. But I looked up high blood pressure, and there's various charts and graphs that show you what is considered to be sort of*

> *dangerously high blood pressure. And so I took my own blood pressure readings [...] and apparently it was serious high blood pressure. So I went back to the doctor, and he said, okay, I'll give you some blood pressure tablets, but... I don't think my doctor is particularly good.*
>
> – **Brian, 54, user**

The low trust in doctors is further corroborated by Brian saying that crowd-sourcing opinions online may be better than relying on the single opinion of a professional. In general, the group of sceptics also thinks that the Internet is a good place to meet others ($\mu = 3.1$; not shown in table):

> *[Online] there's forums and you can cross-reference things a bit better, rather than [depend on] the opinion of one person like the GP or whatever. Depending on how important it is, you can drill further and further, so if you wanted to know more you can. So that's what we do at the moment.*
>
> – **Brian, 54, user**

The group of sceptics tends to think that the Internet has positive effects on their health, and may potentially lower the need for health service use, particularly for minor ailments:

> *Probably if I didn't have the information via the Internet, maybe I would need to go and see the doctor more often. But now as there are other ways of getting the information or, reassuring yourself or whatever, that the doctor is kind of less necessary for minor things anyway. In fact, even if I'd think it would be something more serious, I still would actually look at it on the Internet and find out more about it.*
>
> – **Brian, 54, user**

Other sceptics, such as Barbara or Ethan, concurred with this, saying that they would always go to the Internet for a first or second opinion on virtually all their health concerns, and then only see a specialist for more serious problems or if they needed medication.

4.4.5 The worriers

The worriers tend to find it more difficult to interpret health information online; as a result, health information online tends to have a negative effect on their perceived health and may increase health service use in some cases. People who fall into this category describe health information online as very frightening, and although the group of worriers has the highest frequency of health information-seeking of all groups ($\mu = 1.8$), they report a critical attitude and a stronger (reported) reliance on official websites.

Helen, who is 43 years old, is an example of this group, and is representative in terms of age ($\mu = 41.9$), but with higher education (level 4, $\mu = 2.3$) and socioeconomic status (NS-SeC 1, $\mu = 2.3$, Table 13) than the average in this group. She works as a designer, lives in the borough of Hammersmith and Fulham with her children and suffers from epilepsy. Helen looks up health information relatively frequently, and – like the learners – exhibits interest in reading about health:

> *I can never remember what I clicked on to know whether it's relevant, but it's not necessarily about epilepsy, it's other things. I can spend ages on it, going on things that aren't relevant to me, but I can also really forget most information.*
>
> **– Helen, 43, user**

However, while Helen does look up a lot of health information, she says that reading health information makes her feel very afraid. She also highlights the importance of a supportive environment (see section 6.2), which she considers a necessary condition for looking up health information online:

> *When I look on a website, or when I decide to go on a website, I am afraid of it, I'm afraid of what I might find. And I think if you're on your own and look at a website, and find something really bad, [it's] really dangerous. [...] I think you have to be really in a supportive environment to be able to look up on the Internet some things, it can be really dangerous.*
>
> **– Helen, 43, user**

The group of worriers is also characterised by a comparatively low self-efficacy ($\mu = 2.8$, Figure 13) or external locus of control. While denial of having a health condition is no direct evidence of this, Helen's decision to ignore certain developments shows a certain disbelief in being able to influence things:

> *For many years I didn't make contact, or didn't use the epilepsy website when I was younger because it was too close. [...] I was in denial of epilepsy [...] at one point I wasn't taking my medication because I thought maybe I'd got better... I was proved wrong and had a fit. [Both of my children] will possibly have epilepsy. I haven't brought it up with my doctors. Because I think I'm afraid to do it. And I haven't looked on websites because it's very personal, and I think each person is different.*
>
> **– Helen, 43, user**

At the same time, Helen exhibits high trust in her doctors. Even though she would look up health information, she strongly emphasised the importance of working with medical professionals rather than going to the Internet for looking up health information:

> *Yes, to be informed, and sort of keep in touch with... not necessarily new medications because I don't want to sort of look that up. Well, I can look it up, but I don't take it completely seriously because I work with my specialist and my doctors about medication. [...] It's really important that you work with your doctors and your specialist and not go on the website, because it could really make it worse, you're always looking for the worst information sometimes when you use the Internet.*
>
> **– Helen, 43, user**

Helen also provided an instance in which she insisted on getting an appointment based on something she had read online, and concluded that she had unnecessarily used health services in this case:

> *I've looked up on the website about lumps in the breast and how it's important to have a mammogram. And the doctor said, "Let's see what it's like in two or three weeks, and then we'll make an appointment". [...] And I went "No, this is what we're going to do, I've seen it on the website, you need to be seen as soon as possible, so this is what's going to happen". And so that's what happened. [...] In the meantime the lump had reduced, well not reduced, I couldn't find it anymore. But I went anyway just to have a check-over.*
>
> **– Helen, 43, user**

In general, many worriers presented a strong normative sense of the Internet not being the right place to look up health information, and having been advised not to use the Internet for health information:

> *I do look up health problems, but I only really look up official websites. [...] I've always been told by people not to look up healthcare, if you don't know what you're looking for [...] because you always see the horror stories rather than the positive points.*
>
> **– Helen, 43, user**

For that reason, she also said she would find it helpful to get specific recommendations from her doctor about which websites to use:

> *It would be good if the doctors actually suggested the website to look at, my doctors never say "Oh, look at this website" [...] or "Can I send you an email on that area", they never suggest that. That would be a good idea. Top tip!*
>
> **– Helen, 43, user**

Although Helen admitted looking up pictures of overweight people to feel better about her own weight (see p. 169), the worriers generally stated a strong preference for official information, as opposed to less trustworthy user-generated content:

> *Sometimes you can click on that [illness] and then it will bring up a site that will be more relevant to you. [...] But it would be helpful if it was images, that you actually had the name of the company on that image, or the name of the government body, so you could see it was official, not some person photographing their leg or something.*
>
> – **Helen, 43, user**

Despite the strong normative view of "you should not look up health information online," Helen may be looking up more than she reported in the interview. However, the low reliance on user-generated health information online is in line with other worriers such as Wynne and Anthony, who highlighted that anyone could write anything online, so that they would rather see a doctor right away. In general, the worriers exhibited a strong tendency to rely on professional medical help, as reflected in the high trust in doctors ($\mu = 4.1$).

4.4.6 The delegators

The fifth group is composed of non-users and users who value the Internet as an information source, but do not necessarily want or are able to use the Internet themselves. As a consequence, they often appreciated the comfort of not having to look up information themselves. Compared to the adigital group, they have a higher technology attitude ($\mu = 3.1$), higher learning attitude ($\mu = 3.6$), generally higher Internet skills ($\mu = 3.4$), and higher online enjoyment ($\mu = 3.0$), but a similarly low interest ($\mu = 1.7$) to the adigital type (Figure 13).

This subsection revisits the example of 75-year-old Kathleen from the opening paragraphs of this chapter. At the time of the interview, she (as a non-user) had just obtained information from the local library in order to decide on a therapy:

> *I've got a bad back, so I've been to the hospital, to the spinal unit to have it assessed, and they say that I've got to have some therapy. They wanted me to find out about [osteopathy or physiotherapy], and then make up my mind about what I wanted. So I had to go down to the library and go ask them if they would do me a printout for that, so that I could read about it and see what they do. I don't know whether I really do need to have a computer, because anything like that I can find out.*
>
> – **Kathleen, 75, non-user**

As with the networks of non-users described by Norman (p. 111), Kathleen also has a friend in London who she can ask for information. Her quote reflects both the skill dimension (*"they're very good at that"*) and the device dimension (*"a friend [...] who's got a computer"*) for obtaining information from the Internet:

> *I really just need it for information, so if I need anything, I can go and get it from the library, which I do. They're very good at that. [...] That's given me what I wanted. If I really do want more, I've got a friend in London who's got a computer, and she would always look up anything for me, she'll then phone me back and tell me or she'll send it down to me.*
>
> **– Kathleen, 75, non-user**

Finally, Kathleen also actively reads health information in the local paper, which is a more accessible medium for her, showing her interest in health and her feelings of responsibility for her wellbeing. Though it is not common practice among everyone in this group, Kathleen would also often take information to the doctor:

> *There's always quite a bit in the Daily Mail on Tuesdays [...] [and] it's often those that I take into the doctor. [...] He is always quite amused when I come in and he says "what have you got there for me today?" but he will discuss anything like that with me. So yes, I do take things in, and say I found this or there was that in the paper, and, what do you think, and yes, he will give me his opinion.*
>
> **– Kathleen, 75, non-user**

While the delegators had moderate trust in doctors ($\mu = 3.6$, Figure 13), the unusually high standard deviation ($\sigma = 1.2$) shows the high discrepancy within this group. Kathleen placed high trust in her doctor, and – illustrating again the delegation aspect – particularly valued the fact that he would look up information for her when she took in information from the newspaper to the consultation:

> *It was in the paper that they were doing this particular thing at St Mary's hospital in Paddington, and so I took it in [...] and he said "I'll google that when I get home" [...]. Then he did refer me and I went there. [...] That came about from just the paper, so he will do things like that for me, but he will look into it as well, he doesn't just do it like from what it says in the paper [...] but takes it home and googles it on his computer.*
>
> **– Kathleen, 75, non-user**

For the Internet specifically, Kathleen said that it was not her first port of call to go online. Consequently, the delegators would rather not use the Internet in order to make decisions about whether to see a doctor or not, but rather to learn more about a condition, especially when prompted by professionals to look something up:

> *It was the lady that did the assessment on my back, she said I'd got to find out about the osteopath and so, that's why [...] I went to the library to find out what they actually did and get the printout for that. I normally do look up whatever I got wrong with me, I do like to know.*
>
> **– Kathleen, 75, non-user**

Kathleen's statement also indicates that health information would not worry her. As will be discussed in section 6.2 for the social environment, this can partially be attributed to some level of filtering by those who go online for her. Kathleen's statement about *"using the NHS, not the Internet"* reflects that the responsibility for where and how to look for information is also delegated.

4.4.7 The adigitals

The last group, the adigitals, comprises a large number of non-users, but also users, who do not see the Internet as a useful information tool and presented relatively strong views on the low suitability of the Internet for healthcare.

Charlotte from Cornwall is 78 years old ($\mu = 66.5$, Table 13) and used to be a sales representative covering large areas in England. For this reason, she has a relatively high socioeconomic status (NS-SeC 2, $\mu = 2.7$), but reflects the low educational qualifications of this type (no qualification, $\mu = 0.6$, Table 13). She does not use the Internet, and generally does not like to work with technology, partly for health-related reasons:

> *It goes back an awful long time, I suppose. My fingers have never worked in a way that I can use a keyboard of any sort, piano, computer keyboard, I can't separate them. So I lost patience. [...] Anything to do with office work I absolutely hate. I refused promotion at work, [...] I said no – because it was too much paperwork.*
>
> **– Charlotte, 78, non-user**

For health information, Charlotte would ask the doctor, but most importantly thought that she knew what was best for herself without any additional information. This argument came up in several interviews, particularly with interviewees of the adigital type:

> *I usually call the doctor. [But] I think in most ways you know your own body. [...] [I follow] just my own instincts. And I found out that they never let me down, fortunately.*
>
> **– Charlotte, 78, non-user**

As with the delegators, the interviewees in the group of adigitals did not show a consistent picture in terms of trust in medical professionals ($\sigma = 1.0$, Figure 13). Some, particularly older people, expressed high trust in the medical profession. Others, such as Charlotte, did not place a high value on her doctor's opinion, again highlighting the importance of the "listening to your own body" argument:

> *[I trust doctors, but] only if I tell them what's wrong with me. [...] There are too many cases where they just accept the wages and don't do the work. [...] You have to demand attention. That's the only way to do it. [With my] pulmonary embolism, they wanted me to take paracetamol and get away with it. I knew it wasn't [my] heart, I had no feelings it was a heart thing at all, it was just something I didn't understand. But I wanted investigating before I started disguising [it] with pain killers. If it was up to them, I'd probably be dead now.*
>
> **– Charlotte, 78, non-user**

The group of adigitals generally also expressed concerns about how people use the Internet for health information. As discussed in the literature review, to some extent the hypochondriac is always someone else (Dillon, 2009), but the group of adigitals tended to emphasise this aspect with Internet users:

> *But it seems there's a particular type of person who does it, who loves being ill. [...] They'll worry themselves into goodness knows what and they do the same on the computer – as soon as she sneezes she looks it up on the computer. What's the sneeze, what's this, what's that? For god's sake, leave us alone.*
>
> **– Charlotte, 78, non-user**

While mostly older non-users fall into this type (see Figure 13), it also contains younger individuals such as Sarah (30 years old), who had strong views on how healthcare is becoming increasingly impersonal and should remain offline.

4.5 Effects of Internet use across types

The cases of Miriam, Hugh, Brian, Helen, Kathleen and Charlotte allow for more general observations about the Internet's effects on perceived health and health service use. Although the specific circumstances of the situation (see subsection 5.1.1) and the kind of long-term health condition (see subsection 6.1.4) may play a role, the Internet's effects may be conceptualised based on this typology, since – in Ian's words when he was asked about the effects of health information online – *"it depends what sort of character you are before you go online."*

To this end, this section replicates two analyses from section 4.2 divided by the six types of health information-seekers. Table 15 shows the regression on the two outcomes for users and non-users for each of the six types relative to adigital non-users as the omitted category. It indicates that there are differences between different types of users, but also between non-users, showing the importance of prior attitudes pertinent to each group. It also confirms again the validity of the simulated

dataset by showing that the results obtained by appending both health outcomes of perceived health and health service use to OxIS directly (see subsection 3.4.3) are relatively similar in general. To allow for further comparisons, all other OxIS-based regressions are presented alongside the SM data in Appendix XI (p. 334).

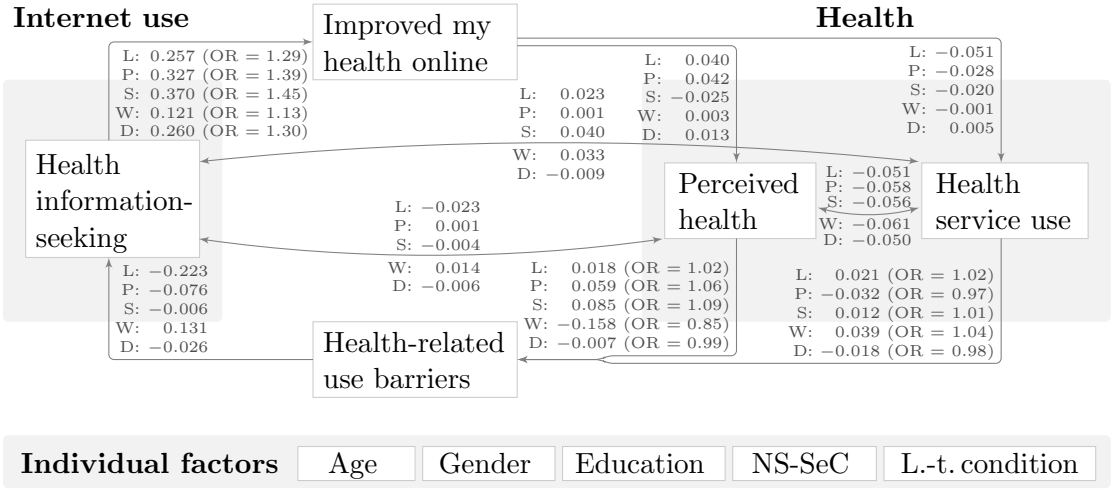
Table 15: Regression with typology of health information seekers

	Perceived health		**Health service use**	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
User of learner type	0.073	0.068 **	−0.043	−0.028
User of pragmatist type	0.047	0.030	−0.031	−0.035
User of sceptic type	0.038	0.019	−0.023	−0.019
User of worrier type	0.041	0.028	−0.011	−0.000
User of delegator type	0.027	0.012	−0.009	−0.007
User of adigital type	0.016	−0.014	−0.010	−0.017
Non-user of learner type	0.014	0.026 *	−0.000	−0.017
Non-user of pragmatist type	−0.005	−0.014	0.009	−0.004
Non-user of sceptic type	0.005	0.025 *	−0.002	−0.047 *
Non-user of worrier type	−0.007	0.028 *	0.017	−0.002
Non-user of delegator type	0.023	0.018	−0.006	0.025
Non-user of adigital type	–	–	–	–
Age	−0.208	−0.161 ***	0.433	0.377 ***
Gender	0.096	0.130 ***	0.148	0.099 ***
Education	0.306	0.306 ***	−0.032	−0.020
NS-SeC	−0.086	−0.012	0.035	0.014
Long-term health condition	−0.473	−0.487 ***	0.032	0.033
Adjusted R^2	**0.60**	**0.56**	**0.24**	**0.18**
N	**42,989,620**	**2,150**	**42,980,620**	**2,150**
Largest CI	**6**	**4**	**6**	**4**

Note: Significance codes: *** < 0.001; ** < 0.01; * < 0.05.

In addition, Figure 14 suggests that effects based on the usage divide differ across types, reflecting the characteristics described throughout the last section. The adigital type has been left out as the model markedly violated the goodness-of-fit thresholds for SEM (see subsection 3.4.3). Both Table 15 and Figure 14 provide the basis for the following two subsections, which present how perceived health and health service use are affected for each of the types. This section bases these statements on the quantitative data and – in contrast to the last section – on other interviewees' voices as well.

FIGURE 14: SEM based on usage divide – split up by typology



NOTE: Paths from the individual factors *(age, gender, education, NS-SeC, long-term condition)* to all variables in the model omitted in figure for clarity. L = Learners, P = Pragmatists, S = Sceptics, W = Worriers, and D = Delegators.

4.5.1 Perceived health

The relationship between perceived health with general Internet use ($\beta = 0.073$, Table 15) and finding something online that improved health ($\beta = 0.040$, Figure 14) respectively is among the strongest for the **learners**. This is in line with the qualitative findings, with many learners saying that the Internet positively impacted their health perception in various ways (see Chapter 5). Ian also captures the duality between finding factual information, but also feeling normal about an ailment based on reading about it online:

> *Absolutely yes [the Internet affects my health perception]! You're not alone, comes to my mind straight away. Embarrassing questions, there again: you're not alone. So, if somebody's took [the] time to make a website for it, they've obviously seen it before, well it's global, the world's seen it before. It just makes you human.*
>
> **– Ian, 42, learner type**

While the relationship between health-related barriers to use the Internet and health information is strongly negative for learners ($\beta = -0.223$, Figure 14), they were also much less likely to claim that their health kept them from using the Internet: out of the 46 OxIS respondents with a long-term health condition in the group of learners (13% of group, Table 13), only one said that this was the case.

In contrast to the learners, many of the **pragmatists** explained that the Internet would not affect how they perceived their health. They attributed this both to using official websites primarily and evaluating the information properly, but then going to see the doctor relatively quickly (or stopping the search):

> *I think it depends who you are, how you take it [...] because I sort of look at everything [...] before I can say that's a correct statement. [And then] I basically go to the doctors, because we have good doctors.*
>
> – **David, 65, pragmatist type**

> *[On] the NHS website, there's just bullet points and [...] it's never too over-exaggerated. It's always very simple, like it doesn't worry you [...] it would usually say "go to the doctor's" or "it should be fine" basically.*
>
> – **Martha, 34, pragmatist type**

At the same time, the pragmatists agreed that the Internet was very useful for health purposes, as it allowed them to seek medical attention when necessary, thereby saving time and effort, and be prepared for the consultation with basic information. Perhaps this explains the positive relationship between perceived health and Internet use ($\beta = 0.047$, Table 15) as well as Internet-based health benefits ($\beta = 0.042$, Figure 14) in the quantitative data.

The medical **sceptics** primarily concurred with the learners that the Internet had positive effects on their health, which is reflected in the quantitative data ($\beta = 0.038$, Table 15). However, while the relationship between health information-seeking and finding something that improved health was reasonably strong ($\beta = 0.370$, Figure 14) – in line with the interview findings – the relationship between finding something online and perceived health is negative ($\beta = -0.025$, Figure 14). This may indicate that online information is not as beneficial as the sceptics might think, or that sceptics generally rate their health lower. The latter, however, could not be substantiated in the interviews, nor was it reflected in the low proportion of long-term health conditions in the group (10% of the group, Table 13).

The majority of the **worriers** explained that health information-seeking in fact increased their concerns. While this is not reflected for their general use ($\beta = 0.041$, Table 15), it is salient that worriers' health information-seeking is only weakly linked to finding something online ($\beta = 0.121$, Figure 14) compared to the other groups ($\beta > 0.25$); in addition, the relationship between finding something online and perceived health is nearly non-existent ($\beta = 0.003$, Figure 14). Anthony explains that, as a consequence of the increased concerns, he has largely stopped looking at health information and goes to see a doctor straight away:

> *I find it a bit too frightening. If you got a little spot on your hands [...] I think the tendency is that before you finish looking on the Internet, you think you might have your hand amputated. [...] It brings up everything, and you end up thinking "oh dear, could I have that" or "could that be...". I think it can be quite counterproductive, quite dangerous.*
>
> – **Anthony, 58, worrier type**

Interestingly, in contrast to all the other groups, the relationship between health affecting Internet use and health information-seeking is strongly positive ($\beta = 0.131$, Figure 14), implying that many individuals who agree that health keeps them from using the Internet would nevertheless keep looking, which may partially explain why the worriers have the highest level of health information-seeking ($\mu = 1.8$, Table 13) despite also having slightly higher levels of those with a long-term condition (17% of group, Table 13) compared to the other four groups which largely consist of users.

None of the **delegators** thought that health information online affected how they judged their health, but said that it may have health benefits, as became evident in the case of Kathleen. At the same time, she argued that she would not worry about the information. This might be attributed to her outsourcing the process of obtaining information entirely, thereby trusting the skills of other individuals and avoiding the need to sift through various online sources. Mary, an example of a user in the delegator group, would rather ask her family to look things up for her, due both to not always having Internet access for financial reasons, and particularly to having trouble finding the right sources:

> *Whenever anything has been wrong with my son, other people check it up for me. [My family] would look it up, and then get back to me, and let me know what it sounds like. [...] When I read things, I take it as the worst thing ever [...] I'm like "oh no, that's bad, that's bad", and then I'm sitting there panicking. [...] So it's just not right for me to [look it up online].*
>
> – **Mary, 29, delegator type**

In support of this, for the delegator type, the quantitative data shows how the relationship for general use ($\beta = 0.027$, Table 15) and health-related use ($\beta = 0.013$, Figure 14) is moderate. It is worth highlighting that the relationship is relatively similar for users and non-users ($\beta = 0.027$ and 0.023, Table 15), potentially supporting the delegation aspect salient in this group.

For the **adigital** users, there is only a weakly positive relationship between Internet use and perceived health ($\beta = 0.016$, Table 15). However, no interviewee of the adigital type attributed any health benefits to the Internet, and the strength of

the relationship corresponds to those of non-users for the learner or delegator type. At the same time, the positive relationship between being a non-user of these two types and higher levels of perceived health (with adigital non-users being the omitted category) supports the finding that factors antecedent to health and Internet use may influence perceived health beyond actual use.

4.5.2 Health service use

With respect to health service use, the majority of **learners** and **pragmatists** suggested that online health information helped them to improve the appropriateness of health services in some cases, which is reflected in the quantitative data for general Internet use ($\beta = -0.043$ and $\beta = -0.031$, Table 15) and health-related Internet use ($\beta = -0.051$ and $\beta = -0.028$, Figure 14). This was true for acute problems in which they needed to take a decision, but also for long-term conditions because the Internet helped to improve one's judgement over time, as Miriam explained (p. 131). Michael, as a learner without a long-term condition, explains how the Internet helps him to reduce health service use by searching for health information instead:

> *You can get a lot from the Internet, you can pretty much care [for] yourself on anything. [...] [The doctor is] the last resort. I'm not kind of a fan of going to the doctors, I feel like I'm wasting their time. It just depends really, if I really feel like I need to go [then I go] but otherwise I really just search online, and keep searching.*
>
> **– Michael, 22, learner type**

The learners particularly highlighted the fact that they might go online to find out more about a condition after the consultation, but – reflecting their learner characteristics – in an attempt to understand more about the condition and side effects of medication, rather than to check on the doctor:

> *If [...] the doctor says "it's this and this [...] and go away", I wouldn't then go onto the Internet and go "is what he's saying right?" I wouldn't do that. I just trust that they know what they're doing – they're medically trained. However, [...] if they give you a drug, and you want to know a bit more about the side effects because you experience something, I might then google that [...] just to understand a bit more about it.*
>
> **– Danielle, 30, learner type**

This is in contrast to the medical **sceptics**. While the users in this group unanimously agreed that the Internet helped them to determine their need to see someone ($\beta = -0.023$, Table 15; $\beta = -0.020$, Figure 14), this group showed a

general tendency to *"not really go to the doctor's if I can help it [because] what's the point?"*, as Barbara argued. Similarly, Ethan also used health information online to replace visits to a physician (Dwyer and Liu, 2013):

> *GPs they haven't really got time for you nowadays. I, as a rule, try to keep away as far as possible. [...] Everything has become so mercenary. [...] They get paid a little more for each patient by getting [them] to have a particular test, and you think to yourself: actually, it's not really me and my condition they are interested in, it's just a way of their pay package.*
>
> – **Ethan, 69, sceptic type**

While this generally reduces health service use, Brian's story about going back to the doctor to prove that he actually had a health problem (p. 136) shows that some information may lead in fact to increased health service use (with the appropriateness of visits not being taken into account). Perhaps this explains why the link to health service use is not as strong as for the learners and the pragmatists.

Beyond Helen's example of increased health service use based on *"I've seen it on the website"* (p. 139), none of the **worriers** recalled specific situations in which they went to see a medical professional based on information read online. However, they agreed that there was a possibility that this might happen, and that the Internet would not necessarily decrease their health service use, particularly due to their view that seeing a medical professional was the right thing to do:

> *People who looked it up on the Internet, they tend to be quite paranoid afterwards [...]. It's too much information, or too much wrong information. So I don't tend to look on it now. [...] I'd go straight to the doctors.*
>
> – **Anthony, 58, worrier type**

This is in line with the qualitative data: while Internet use in general was linked to somewhat lower health service use for the worriers ($\beta = -0.011$, Table 15), finding out about a health-related questions online was barely related to either perceived health ($\beta = 0.003$) or health service use ($\beta = -0.001$, Figure 14). In that sense, the Internet may not lead to more health service use, but it suggests that worriers do not obtain as many benefits from their health-related Internet use.

For the **delegators**, the Internet was not really a decision tool about health service use, partially because going to the doctor was just more convenient than accessing the Internet. This is reflected in the generally weak relationship between Internet use and health service use ($\beta = -0.009$, Table 15; $\beta = 0.005$, Figure 14). Supporting the delegation aspect in this group, it is noteworthy that the differences between users and non-users are small ($\beta = -0.009$ and -0.006, Table 15). Overall,

delegators predominantly saw the Internet as a tool for deeper investigation of certain topics, as Norman explained (p. 111), and which thereby might increase certain elective procedures, as in Kathleen's example (p. 141).

Similarly, Internet use barely had an effect on health service use for the **adigitals** ($\beta = -0.010$, Table 15), but rather for reasons associated with the quality and usefulness of information online. In addition to Charlotte's *"[I follow] just my own instincts"* (p. 142) argument, many of the adigitals highlighted the importance of traditional offline services, and voiced their mistrust in information online:

> *[If I have a health question] I'll ask somebody who knows, not the Internet! I mean really, that's the way I understand it – anybody could put any rubbish on there and you would think it was kosher, and I'm not interested in that. That... and in fact I probably wouldn't know how to do this.*
>
> **– Shirley, 71, adigital type**

Finally, some statements indicated that the Internet would not necessarily change the quantitative metrics of health service use and perceived health, but rather the nature of their consultations, as is clear from Brian's statement (p. 136) about demanding attention, and Kathleen's example of initiating discussions during the appointment (p. 141). In addition, there are differences in the *way* in which the Internet influences health between these types, which are also reflected in the residual relationship between health information-seeking and the two health outcomes not attributable to any of the two causal paths ($-0.023 \leq \beta \leq 0.033$, Figure 14).

Consequently, the next chapter draws primarily on the qualitative data, giving full particulars on the ways in which the Internet affects perceived health and health service use. While this next analysis focuses particularly on Internet users, and the mechanisms in Chapter 5 are not type-specific, the established typology is used as a lens for understanding these mechanisms where applicable.

Results 5

Mechanisms of Internet use and health

Ian moved to an urban area on the Isle of Wight in 2013, after growing up and establishing himself with his wife and his children in a rural village on the island. The move, at the age of 40, entailed various changes in his everyday life, such as a switch to a new GP and a dramatic improvement in mobile phone network coverage and Internet speed. Ian (classified as the learner type) says that in the past, he would never have thought of streaming documentaries or quickly looking something up, whereas now both of these activities opened up *"a whole new world with all that knowledge"* for him.

Throughout the interview, Ian provided numerous examples of how the Internet influenced health-related aspects of his life in various ways. Later that week, he would go to the local hospital to receive *"the yes or no for a minor op"*, which resulted from an online search on what Ian only referred to as his *"little problem downstairs"* and had kept secret for the last nearly 30 years:

> *[The consultant] said I've done the right thing. She asked how long have I had it, I said "since I was like 14, 15". [...] And she said: "How come you've put up with it for so many years?" And I said, "Well, you do, it's one of these things that you won't bring to the doctor's attention" [...] And because of the Internet [I realised] I've got rights to deal with it. [...] It was the Internet that made me feel as though I am normal, I am entitled as a human to go and question this situation that is embarrassing. But prior the Internet, prior to anything, I would have left it.*
>
> **– Ian, 42**

Beyond feeling legitimised to act, Ian appreciated the personal access to information before explaining himself to anyone else. He also portrays the Internet as a knowledgeable, caring person who guides the way (Pols and Moser, 2009):

> *Once again, good old Internet, showed me, talked me through [it] [...] am I right checking myself, am I doing this right [...] it's a very personal thing, no one can see what I'm doing, no one can see what I'm up to, and personally I feel confident – I don't have to explain it to the doctors!*
>
> **– Ian, 42**

The information picked up online not only encouraged him to make an appointment, but also reconfigured the consultation itself by providing him with keywords:

> *I was sort of like, thank God for that, now I can make a GP appointment, [and] know what to say because I had references from the Internet. [...] I went in, made myself an appointment, used keywords from what I had picked up from the Internet, [but] never mentioned the Internet.*
>
> **– Ian, 42**

Moreover, Ian explained that the *"computer"* would also learn more about him, including his obsessive-compulsive disorder (OCD). Weighing the benefits and drawbacks, however, he generally valued the associated convenience:

> *The thing that I find a little bit nosy with the computer is it's reading me, it's educating my habits, my trends... or mental health. But [...] I'd say it's a good thing, because I don't have to go [...] into the search, where it's learning of me – I'm kind of being a bit lazy.*
>
> **– Ian, 42**

However, Ian clearly distinguished between what the computer knew about his mental health, and what other individuals did, for example those on an anonymous, hobby-related network. He saw the Internet as a stigma-free space to feel normal:

> *I kept that to myself [...] I didn't want to go on the Internet as a target. I'm aware of the stigma for having mental health [...]; there's people that I've known all my life that doesn't know a dribble of what I endure every day. Online, it's like, I join as a normal character without the mental health. So I'm going to keep that out of everything I do, every site.*
>
> **– Ian, 42**

Ian's example shows the various ways in which the Internet may influence health perceptions and health service use, as addressed by RQ_2 in this research. In particular, it highlights four interrelated aspects that will be detailed in this chapter. First,

the Internet may both increase and decrease health outcomes (Berger et al., 2005; Suziedelyte, 2012): while Chapter 4 provided examples of how individuals reduced visits to the doctor based on information online, Ian was encouraged to finally seek support for his embarrassing condition. Second, health information online may not just affect the frequency of health service use, but also the consultation and relationship dynamics itself (Murray et al., 2003), for example through the learning of appropriate words to use. Third, while health information online may be passed from one individual to the other (as printouts, for example), some aspects relate to actual use and the Internet's mediation affordance, such as providing personal access without having to discuss a condition with anyone else; this is particularly applicable for embarrassing conditions. Finally, beyond the Internet's content affordance, there are aspects around mediation and connections to others, for example with the Internet as space to be free of stigma, which may influence an individual's coping ability and positively influence health perception.

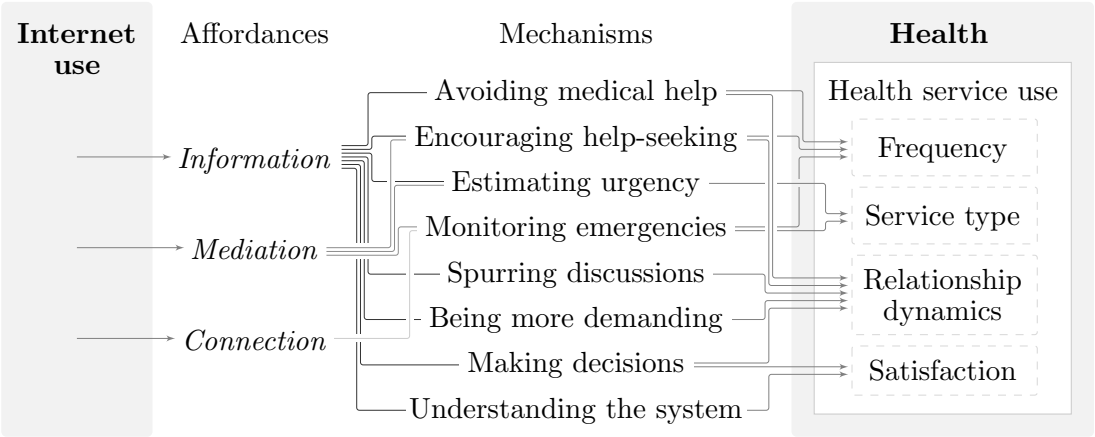
While all of these aspects are closely interwoven, the next section (5.1) will present the effects of Internet use on health service use, followed by the results for health perception (section 5.2). Although the quantitative data can only partially capture the richness of this relationship, section 5.3 integrates some of the aspects discovered into quantitative models to measure the emerging relationships on a larger scale. Where applicable, interviewees are presented alongside their type, using the typology developed in Chapter 4, as some mechanisms between Internet use and health are more typical of certain types, or show different manifestations across types, although the mechanisms presented are type-neutral in principle.

5.1 Effects on health service use

Based on the qualitative interviews, the Internet has shaped health service use in four major ways. It has affected whether health services are sought at all, and what kind of service is sought based on the urgency required. In addition, online activities shape the interaction and relationship dynamics with health professionals, but also the satisfaction with health services and the wider health system.

Figure 15 shows an overview of eight observed mechanisms for each of these facets of health service use, all of which are detailed throughout this section. Based on the patient-centred methods used in this research, it goes further and evaluates whether seeing a specialist was necessary or not. The focus is rather on how the Internet influenced the decision to seek professional help or not.

Figure 15: Effects on health service use based on qualitative research



5.1.1 Avoiding medical help

Many interviewees reported they have tried to use the Internet to avoid bothering the doctor or to save their own time (see subsection 4.3.1), also showing how health information-seeking is embedded in everyday health practices (Kivits, 2009):

> *It's good to check, because you don't always want to [...] waste the doctor's time. It depends on what it was really, but if you can get away with looking it up online... sometimes it's just [...] a niggle in the back of your mind, and you're like "I'm not going to the doctor's [...] I'll google it instead."*
>
> – **Isabelle, 38**

The Internet helped individuals to evaluate situations, or provided advice for relieving ailments. However, whether this worked out depended on the circumstances. First of all, it depended on the kind of health problem:

> *If it's something I never had before, then I [...] just go and sort of see what [they say]. If you've gone to a couple of sites, and they all sort of say the same thing, then you can kind of sort it out by those symptoms.*
>
> – **Margaret, 57**

While many individuals viewed the Internet as a resource for first-time problems, Miriam also re-read a checklist of stroke symptoms from time to time (after already having a stroke), which helped her to avoid bothering the doctor in vain:

> *I've read about [stroke symptoms] a few times since it happened as a refresher, and [...] know, none of this happened, so I'm alright. Rather than go to the doctor's and say, this has happened and blah blah blah, I would go on there and re-read through it. So I find that really helpful.*
>
> – **Miriam, 59**

Wanda adds a second decisive factor around the severity of the health problem at hand. In her case, there was a relatively low risk associated with her insect bites:

> *[I] saw photographs similar, and sort of hand and heart made the decision not to follow it up and do anything else about it. Which was the right decision, because it did sort itself out in the end.*
>
> – **Wanda, 58**

Wanda's quote again highlights the importance of using images to evaluate conditions, which were also reflected in Jeremy Hunt's statement on his child's skin rash (p. 18). In a similar case with higher severity, Matthew says that although the problem was in fact solved through the Internet, he nevertheless sought medical help due to the potential consequences of misdiagnosed meningitis:

> *[We learned online] if you push it with a glass, and it doesn't disappear, then you should go, it's quite bad. A normal rash disappears when you push it with a glass. [...] [The Internet is] just quite useful to know what [our son] has. In fact we still went to the doctor's [...] you just worry about meningitis, the brain thing.*
>
> – **Matthew, 47**

The difficulty of identifying symptoms properly relates to a third decisive factor for effects on health service use, which centres around how clearly symptoms could be described, and how broad the spectrum of search results was:

> *I looked about having back pain, but there's such a lot of information that I really didn't know what I was looking for. So I thought, no, I'm going to see a doctor and let them sort it out. [...] There was a lot of information, and lower back pain could be a symptom of a hundred different illnesses.*
>
> – **Carol, 61**

In this context, the pragmatists and the worriers in particular highlighted the importance of listening to the expert:

> *Sometimes [the Internet] can just ease your mind at first. [But then] you're like "This can't be right," and you do need the expert [...] to tell you if it's this or not. Best to go to the GP – someone to check you out.*
>
> – **Brenda, 40**

This stands in contrast to Ethan, who thought that *"it's not really me and my condition [doctors] are interested in, it's just a way of their pay package."* As a sceptic, he would try to find an answer to his question with the help of medically knowledgeable friends. This points to a fourth decisive factor in terms of medical expertise in the individual's network:

> *[My advice is:] Use Safari in the initial stages, try and inform yourself a little more, and then try and talk to people that you know to confirm that that's the way forward. [...] I would talk to one of several medical friends that I have, which was how things used to be, and they're still pretty reliable. [...] Because GPs, they haven't really got time for you nowadays. I, as a rule, try to keep away as far as possible.*
>
> – **Ethan, 69**

Similarly, Michael would talk to a known person first, though, in contrast to Ethan and reflecting his learner-type characteristics, he wanted to find out as much as possible himself without bothering medical professionals:

> *I look up a lot when I feel a bit ill. [...] I look up things, and if it doesn't satisfy me, I just go deeper and deeper. And if that still doesn't satisfy me, I just go for general advice of a person I can relate to. [...] To be honest, I do look up a lot, but I find it a lot easier talking to someone than just reading the facts or quotes or whatever.*
>
> – **Michael, 22**

These examples underline the relevance of the social environment (discussed in section 6.2). Other important influences for understanding the effects on health service use, such as education, are analysed in more detail in section 6.1.

5.1.2 Encouraging service use

Health information online may also increase health service use, as "many patients are too embarrassed to ask questions they think foolish or highly personal" (Anderson et al., 2003, p. 73). The example of Ian's health condition, which he concealed due to feeling embarrassed (p. 151), showed how the Internet may help people to feel normal, legitimise help-seeking, and provide keywords for the consultation.

Beyond stigmatised conditions, finding out about treatments online may actively support health service use, particularly for elective procedures. Brian, as a sceptic, found out everything about eye surgery online, and then intended to contact a professional to get the procedure done. Reading others' stories also encouraged him to pursue his plans:

> *I've been looking at [laser eye surgery], just trying to find out what people say about it, and would it work for me. [...] Kind of reassuring when you see there's more people who have it done and had to deal with the same kind of anxieties that I would have, about putting yourself in someone else's hands. Everyone is squeamish with their eyes.*
>
> – **Brian, 54**

Ahmad, who had NHS coverage for emergencies, but had to pay for all other treatments, said that reading about his sports injuries online created a desire to see a specialist to fix his ailments more permanently:

> *Because of the information of the Internet, it makes me want to see the doctor, because I have a problem. It's impossible to skip my injury all the time, so I need to find something to cure my injury. [...] I want to go and see the doctor because I know they can help me with my problem. But the problem is I'm just scared how much they cost me.*
>
> **– Ahmad, 23**

Particularly for the pragmatists, Internet use may increase health service use if online resources specifically recommend seeking professional help. For Martha, the Internet has facilitated the decision to call a doctor:

> *[My daughter] woke up, wiggled down and fell off the sofa before I could catch her. So she had a big bump on her head [...] So obviously I looked that up online, and in fact, because it said follow it up, I phoned the doctor. [...] She was fine. So that was really helpful.*
>
> **– Martha, 34**

Then again, the Internet may also increase health service use based on reading about extreme cases online, or exacerbating worries about potential conditions, as when Helen insisted on getting an appointment based on reading something online (p. 139). While health information online can be frightening, experience with the Internet and, in this case, reading about health online, can help users to become acquainted with the Internet, with experience generally being an important influence on Internet-based outcomes (Haythornthwaite and Wellman, 2002):

> *When I first started using [the Internet], it panicked me, and I could have been a botherer to the NHS. It could exacerbate how you're feeling tenfold, 9 times out of 10 it exacerbated the urge to see a GP, but now that I'm used to it, I find myself not worrying too much about it, because the Internet is a point of reference.*
>
> **– Ian, 42**

Ian's account also highlights the interrelated nature of both health outcomes in this research, as the Internet might affect one's health perception (see section 5.2), thereby creating the urge to seek medical help.

5.1.3 Estimating urgency

Beyond effects on frequency of health service use, the Internet may also influence how individuals judge urgency, as reflected in Dwyer and Liu's (2013) finding that

Internet use decreased the likelihood of accident and emergency visits. Several interviewees stated the Internet made them realise that help was not as time-critical as they initially thought, and they valued the specific and actionable advice online:

> *Is it something more serious that I need to get a GP or an ambulance? [...] If it says "nothing to worry about" or "leave it 48 hours," then I just follow the advice that is given.*
>
> – **Danielle, 30**

Information online may in fact be more actionable. For example, the Internet gave Miriam a specific number of hours between having stroke-like symptoms and being at the hospital, as opposed to her doctors' general advice:

> *[The doctors] basically said, if your vision goes blurry, you may be having another stroke. But on the Internet it says "you have to be at the hospital within 4 hours" for them to have the scan and confirm what it is.*
>
> – **Miriam, 59**

Miriam also evaluates the required urgency of dealing with potential side effects online, thereby determining what kind of medical support she would look for:

> *When I first started to get the numbness on my face, I was panicking because I thought I was having another stroke. But [...] on the Internet, it said after having a stroke [...] [it] was quite common to have it on part of the face. So I was more aware that it wasn't this urgent. I still went to the doctors, but I didn't dial the emergency number because I wasn't having another minor stroke. I knew it was part of the thing.*
>
> – **Miriam, 59**

Then again, Helen describes a case in which she did not look up her symptoms online, and hence underestimated the urgency that would have been required:

> *I had my appendix out, and it was the most unhelpful. Maybe if I had looked it up on the website, I could have gone straight away and not waited for an appointment. [...] That was really bad.*
>
> – **Helen, 43**

As these instances show, estimating urgency may be perceived as helpful for acute problems, but also for long-term conditions or returning incidences, as with a stroke.

5.1.4 Monitoring for emergencies

Related to evaluating urgency, the Internet may play an important role in prompting people to seek medical help in emergencies – either through technology itself, or

with the help of others. Kevin, for example, exhibited a very low attitude towards technology, but valued his mobile for emergencies:

> *Before I had the stroke, I hated phones. But now, like, it's good when you need it. [...] When I had a stroke, [my dog] alerted my neighbours. I don't want to hear her bark like that again. [...] [Now] there's lots of people who got my phone number. If something is wrong, I ring them up and [...] they sort it out for me.*
>
> – **Kevin, 40**

Those in particular with a long-term condition, or those who had already survived a major health incidence like a stroke, were often proactively monitored by family members or a nursing home warden. Miriam, for example, talks on the phone and exchanges messages with her daughter, who lives within the same OA:

> *When I didn't answer my phone, sometimes she'd keep trying for about an hour, and then asked "Mum, are you alright?" "I'm fine, I was on the phone." But yes, she texts me, and then she phones every day.*
>
> – **Miriam, 59**

Naturally, monitoring is not necessarily restricted to the Internet. Charlotte, for example, carries around a traditional mobile phone in her bra for emergencies:

> *I have got stuck in the house twice, when I have been on the floor and couldn't get up. [...] We had these things round our neck to give warnings, and last year they decided to start charging 3.95 a week. That was more than the pension increase we got. So I thought what to do and [takes out a mobile] that's how I went away with it. That's always in my bra.*
>
> – **Charlotte, 78**

In contrast, Internet-enabled devices directly linked to medical organisations also help monitor certain conditions. Nathan, who suffers from heart failure, checks his weight on a daily basis to detect early signs of cardiac decompensation, and has his heart rate monitored day and night by the local hospital. He highlights how this helps to reduce the frequency with which he would otherwise be checked up, and ensures medical help in emergencies:

> *I used to have regular visits for monitoring the heart, to make sure it was alright, but now we have a monitor from the doctor that looks after my health, apparently. It's in touch with the hospital, and if anything goes wrong it signals. [...] It's excellent as far as we're concerned, because it looks after the health. And as far as the hospital is concerned, you're not tying up time and beds just to be there and get examined.*
>
> – **Nathan, 90**

In addition, Nathan confirmed that the *"monitor that looks after his health"* also fosters his feeling of being adequately looked after, again highlighting the multi-faceted links between health service use and perceived health.

5.1.5 Spurring discussions in consultations

Beyond the frequency and urgency of health services, several interviewees mentioned that the Internet affected their relationship with doctors and the consultation as such. To some extent, these two are still related. As Margaret deals with minor ailments herself, she feels that her doctor takes her more seriously:

> *I only go if I really need to, because I think when I go, my doctor treats me quite seriously – because he thinks, actually if you have come, there is a good reason for you coming. I'm not just coming because it's somewhere to go for half an hour. So that's the relationship I have with my doctor.*
>
> **– Margaret, 57**

Several interviewees mentioned that the Internet would allow them to ask more questions since they would be more informed in advance of the consultation:

> *On question and answer pages, those questions, you think, I'll ask the doctor that about me. [...] It stays on the back of your mind, it just prompts you to ask questions. Because normally I just go to the doctors, tell him what's wrong with me, and don't ask anything. But now I'm a bit more like "what if this and that".*
>
> **– Miriam, 59**

Similarly, Mary as a delegator who obtained health information from her family, uses the information to hide her worries and appear more knowledgeable:

> *[I use the information my sister looked up for me so] that I can say, "oh, it's not that bad", or "I know what I got to do", "I've got to do this," or "we've got to do this," and he would be like "that's fine". But if I read it was the worst, and then get told it isn't, then I'd be like "oh my God", I'd just be thinking I'm a silly person for thinking the worst.*
>
> **– Mary, 29**

Although taking information to consultations only rarely resulted in negative experiences (Murray et al., 2003), several interviewees would not give away that they looked things up online, but rather use the information tacitly:

> *One thing I wouldn't do is say to the GP, "I've read it online." Because the GP would now look at you thinking you've self-diagnosed yourself and you're now putting the GP at business. [...] The Internet's fine for them,*

> *when the door is shut and they got no patient, but if you're on the other side of the table and you dare mention that you've been online, you're now questioning their skills.*
>
> – **Ian, 42**

However, Wanda reports that even without mentioning the Internet, she got derogatory remarks from the doctor:

> *I mentioned it, and was a bit scoffed at. [...] I didn't say "I looked it up on the Internet," I said "I'd read", and she said, "Oh, you've been on the Internet, have you?" in a sort of derogatory way, a bit dismissive. I mean I appreciate, it's probably difficult for GPs when people go in every five minutes, but [...] I'm not one of those people who sits there and googles this and that and thinks, and "oh, I must run to the doctor's."*
>
> – **Wanda, 58**

Particularly for long-term health conditions, many interviewees used health information very openly, and took advantage of medical professionals as a sounding board. Whitney generally exhibited low trust in doctors, but valued the consultants who were involved in the actual research, bringing in information she found online:

> *Because there is no cure for MS, and they don't know what causes it, people tend to catch on to anything, and sometimes take things that they really shouldn't. [...] I just use [our consultants at a major hospital] as a sounding board really to ask, "what do you think about it?" because they're actually involved in the research. So it's best to ask them as well.*
>
> – **Whitney, 60**

The "doctor as sounding board" idea was also apparent for non-users. Kathleen, for example, took in printouts and newspaper articles to discuss treatments with her doctor (see p. 141), while Natalie showed books or printouts obtained from her partner's sister to her doctor:

> *I say obviously to the doctor, "I found this in the book, what do you think?", and he usually says, "yeah, I agree with what they said" or he says, "oh no, ignore that, follow this way first." Because obviously sometimes the book can be misleading, sometimes, and you're thinking, you just want to check before you follow that information.*
>
> – **Natalie, 31**

The example of the book is a parallel to online health information from official sources, illustrating that health information may shape consultations independently of online health information.

5.1.6 Being more demanding in consultations

Beyond spurring discussions, some interviewees explained that the information they found would help them to be stronger and more demanding in consultations – as they said, in cases where professionals may overlook symptoms or take them as something benign, when they could be an indicator of something more serious:

> *Unless you got serious symptoms [...] your doctor would think it is something smaller at the beginning anyway. So I [...] look it up and [could] be a bit more demanding if I knew it could be a sign of something. Because doctors are so busy that they're like "it's just another cough" or something – I don't think they mean it, it's just how overstretched they are at the moment.*
>
> **– Martha, 34**

In contrast, the sceptics tended to be more openly challenging towards the doctor, such as Brian who took his own blood pressure readings and interpretation tables from online resources (p. 136) to a consultation to convince the doctor of his problem, or Donald, who sent the doctor what he thought was the solution to his Benign Paroxysmal Positional Vertigo (BPPV) after the consultation:

> *I went to the doctor, and she had no idea. She was only a young doctor. So I had a look online and thought, "yeah, that's it." So I printed it off, put it in an envelope and sent it to her.*
>
> **– Donald, 74**

Helen's previous example about the making an appointment for a potential lump in her breast (see p. 139) shows how individuals also gave away the Internet as source to be more demanding, using phrases such as *"I've seen it on the website [...] so this is what's going to happen"* to legitimise actions.

5.1.7 Making decisions about treatments

Several interviewees also found the Internet helped them to make decisions about certain treatments. Partially this was due to being able to read about treatments online, and finding out more about what is being done:

> *I don't like needles and I don't like hospitals, but because I've read it up exactly what they would do, and how long it would take, and all the other things [...] I was fine about going in. It did settle my mind, because I knew exactly what they were going to do.*
>
> **– Miriam, 59**

The sceptics in particular wanted to be involved in treatments, and make sure that they had a say in what is being done. Elizabeth, for example, reports an incident when her sister suffered from gastric cancer, and now was fine despite deciding against the more drastic treatment proposed by the specialist:

> *[Doctors] have to prove what they're telling me, I really do insist on that. Because some doctors just give you medication and say "don't worry about it," and ten years later [...] all of a sudden it's cancer. [...] It's good listening to them, because they are the professionals, but you must look at it with common sense [...] after all you know your body better than the doctor knows it, because it's yours. You see it all the time, they only see it part of the time.*
>
> **– Elizabeth, 55**

In addition, the Internet has helped Danielle to realise what is possible, and therefore set the bar higher for treatment. While her three-year-old son's eczema had significantly improved throughout his life already, Danielle found out online about possibilities involving less medication:

> *It could [have been] a case of "well, just follow people's advice and carry on with the way we are," but now I know that other people are trying natural remedies, I want to try these remedies to see if we can get his [eczema] under control without the medical drugs.*
>
> **– Danielle, 30**

As presented in subsection 4.4.6, online information also helped non-users such as Kathleen to make decisions about physiotherapy versus osteopathy. This was especially the case for non-embarrassing conditions, where the Internet's information affordances contributed most of the value. Of course, not everyone is keen on making decisions (Powell and Boden, 2012). Catherine, for example, decided immediately to have her disintegrated spine operated on, hardly understanding why one might need time to think about it:

> *There's no messing about. I mean there's a couple of people with practically the same problem; the consultant told them what he was going to do, and they were going home to think about it for two days! What's to think about? [...] I was already getting better when they were getting theirs done. So if he said it needed doing, that's it.*
>
> **– Catherine, 77**

In contrast to Catherine's view, several others across all types said it was good to be involved, even though only the sceptics strongly advocated for actually having a say in the treatments and making decisions.

5.1.8 Affecting health service satisfaction

From a broader perspective, the Internet may extend individuals' horizons by bringing in views from other countries, thereby increasing satisfaction with the health system. While many interviewees preferred UK over US websites (see subsection 5.2.1), online documentaries have also made Ian appreciate UK health services:

> *I was watching this documentary online about how people cope without the NHS in America or Canada [...] I found that documentary very disturbing; the laws are different, which reflects on how lucky we are being UK citizens. We're protected a lot more in a lot more ways, so thumbs up to the Internet from my point of view!*
>
> **– Ian, 42**

Many interviewees valued the Internet for service provision (such as booking appointments, obtaining repeat prescriptions, or getting test results). However, the more impersonal nature of healthcare delivery online also created dissatisfaction:

> *You can't phone anymore, you have to do it all online: ordering prescriptions by phone, they've cut making appointments [...] I don't like doing that sort of thing online. [...] And there is no personal touch now, is there? It's all electronically done, you don't speak to anyone [...] it's just straitlaced, take a number and sit down.*
>
> **– Sarah, 31**

From a different point of view, Hugh – in line with other pragmatists – explained that he embraced the efficiency that the Internet may bring to the health system due to the need to save money. In that sense, the Internet positively affected his outlook on how care may be delivered in the future:

> *The truth is that you cannot afford the level of care that people enjoyed in post-war Europe. [...] The service model [of the NHS] is still set in 1950, and perhaps technology allows us to treat people very differently. [..] People are going to have to look after themselves increasingly in society, [...] and obviously the Internet is going to play a core role in that: people selling their services or servicing people through the most efficient ways.*
>
> **– Hugh, 45**

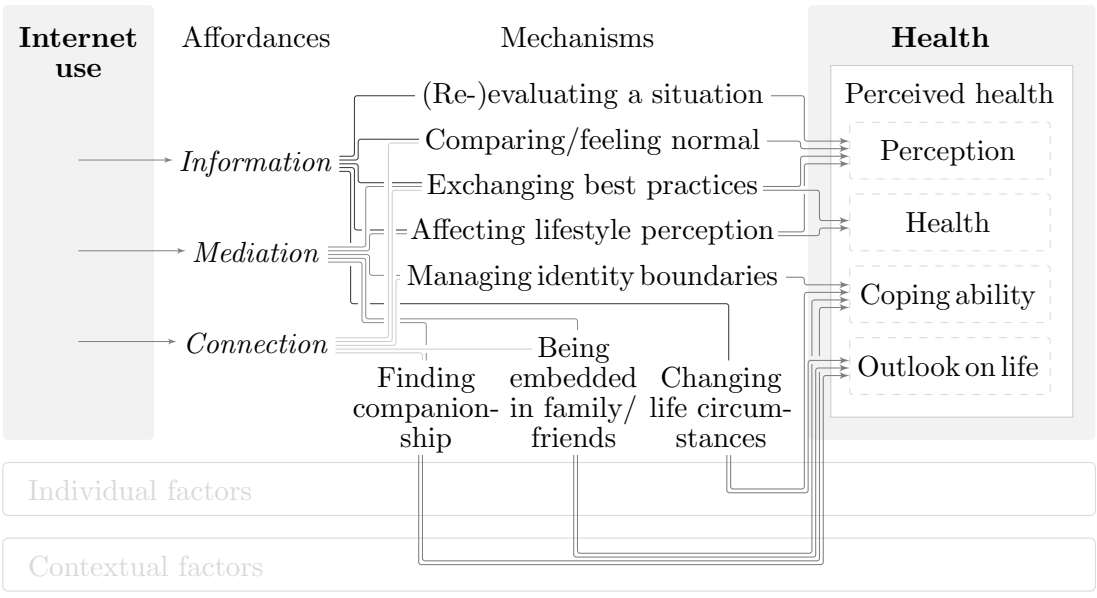
These statements evidence how the Internet affects health service use in tangential ways, and how the Internet shapes the overall view of the health system. This perspective is particularly important for deriving implications for practice and policymaking. Before these are discussed section 7.4, the next section complements the findings so far with a focus on mechanisms for health perception.

5.2 Effects on perceived health

Health perception was influenced by various direct and indirect ways as well: the Internet helped individuals to (re-)evaluate situations, changed their relative perception by allowing them to compare themselves to others, helped them to tap into user-generated advice, affected their lifestyle awareness, and helped them to manage identities and boundaries in their daily lives. With indirect health benefits, the Internet also enabled them to find companionship, be more embedded with family and friends, and improve their life circumstances more generally.

All these affect perceived health in different ways: some may directly influence health, for example through the exchanging of health information, or by promoting a healthier lifestyle. Others rather affect the perception of one's situation, or image of oneself, which is equally important in terms of individuals' self-efficacy to change their circumstances (Bandura, 2004). Several mechanisms work in more indirect ways, for example through individuals' social network or their sociodemographic circumstances, and may affect their coping ability or general outlook on life. Figure 16 systematises these effects, which will be detailed throughout this section.

Figure 16: Effects on perceived health based on qualitative research



5.2.1 Re-evaluating a situation based on information

While evaluating situations already played a role in terms of how individuals made decisions about health service use, it may also help them to (re-)evaluate a situation

and feel reassured, particularly for issues that individuals did not deem important enough to warrant professional help. As with Whitney's aim to avoid worrying *"about things in waiting"* (p. 116), the Internet eased individuals' minds:

> *I think [the Internet] is a bit like a library, you can quickly look up like an answer that you would be worrying about otherwise, and so you can quickly put your mind at ease if you go to the right sites.*
>
> – **Martha, 34**

The addition about the "right sites" reflects Martha's pragmatist preference for official websites. The "easing one's mind" theme also appeared in Carol's reflections on the time when her underactive thyroid was being diagnosed. The Internet helped her to feel better by adjusting her expectations about the time required to find the right medication. In that sense, she re-evaluated her current state of health and it positively affected her outlook:

> *It's reassuring, because you know what to expect. When I first looked [hypothyroidism] up, [...] people were saying that it takes a long time to sort out the dose [...]. And people said that once they get the dosage right, you feel much much better... which is also good to know, that you are going to get better, and that's when you feel better already.*
>
> – **Carol, 61**

This may particularly be true in situations where the patient has genuine concerns. Ian described how he felt very puzzled after seeing a doctor with whom he had several communication difficulties. The Internet helped him to feel relieved after a problematic consultation in the context of his septal defect:

> *This doctor was so kind and so sweet, but he had such a strong foreign accent [...] So I came out and had a bit of a cry to myself and this nurse was like "I'll go and get some notes" [...]; there was keywords, and I just wanted to know what it meant, and then it was, "oh, that's what he's meant, alright, [...] not a problem."*
>
> – **Ian, 42**

However, information online can also negatively influence how individuals perceive their and others' health. The interviewees provided some instances in which the Internet confirmed their fears, or provided them with information that did not help to ease their minds:

> *[It said online that] people can live a real long live with this disease [that my dad has]. And then there were people who said "Well, if you don't get some sort of virus." And I was reading that thinking, "My dad's got a*

> *virus" [...] That really didn't help! [...] I worried more, for no reason. I found out information, which was good, but I found myself worried more about it, so it didn't really help in the long run.*
>
> – **Isabelle, 38**

Beyond increased concerns due to the nature of the information itself, two aspects in particular were associated with unnecessarily arousing negative feelings in particular: exaggerated or unsuitable information; and incidental information found online. As with Isabelle's example, exaggerated information online primarily referred to user-written information, which may overstate circumstances, miss out the particular history of the person, and relate to a wide range of symptoms:

> *They got a past history that you know nothing about, and I think if I explain myself to somebody, I miss out parts of information or exaggerate, not solely the good things, but people exaggerate bad things to make themselves more interesting.*
>
> – **Helen, 43**

> *On blogs, people write "I have this" and "I have that." Every situation is so unique [...] symptoms that I might think of as a sty, someone else might think it's a scratch, so they're seeing it differently in the first place.*
>
> – **Martha, 34**

In addition, the website's origin played a role, as health information on US sites tended to create more concerns for the interviewees (Nettleton et al., 2005):

> *It varies from country to country, so you can get American sites that look at it one way, and British sites that look at it another way. [...] [American sites] go over the top there, I think we are much more measured in England.*
>
> – **David, 65**

Different types viewed these effects differently. In contrast to Helen, a worrier, Martha and David, both pragmatists, said that while they noticed these differences, they were sufficiently *"level-headed"* (p. 120). These differences occurred across types despite the individuals having similar sociodemographic backgrounds (with Helen and Martha both having university degrees). As Ian highlighted (p. 157), experience also buffers negative effects from exaggerated health information online.

Incidental findings occur when individuals look up one thing and receive more information than originally sought. Information found by chance may be problematic when individuals deliberately choose not to look up certain information. For example, when Helen and her mother looked up information about her mother's

melanoma, they decided to not look up certain aspects. While Helen's quote gives away that she in fact looked up the information herself later, she also highlights some "social filtering" that will be discussed in section 6.2:

> *We sort of looked up information about it [...] [but] we didn't type in "what happens?" or "what is the death rate?" or "what is the percentage of people that survive?" We did not look that up, which we could have done. Because actually it was horrific, the percentage...*
>
> – **Helen, 43**

However, learners like Ian and Miriam highlighted that serendipity in their searches helped them to come across topics of interest (Tian and Robinson, 2009):

> *That's the lovely thing about the Internet, you don't have to get really mindful about what you're putting in, just put in a word that you got on your mind [...] it's so good, all the knowledge [that comes up].*
>
> – **Ian, 42**

> *Sometimes people take the wrong things, so if they type one word into Google, and it brings up something completely different, it's connected. [...] But then if somebody says, like it happened to me last week, I've got so and so and so [...] you could explain what you've read.*
>
> – **Miriam, 59**

Miriam's quote also highlights her interest in helping others with the medical knowledge she obtained online; this is a trait that occurred in several learners, which again links to the relevance of the social environment (see section 6.2).

5.2.2 Comparing and feeling normal

Two aspects are related to comparative effects in the context of others: comparing oneself to others, and realising that many are worse off (Miriam: *"it just made me feel like 'stop whining about it and live your life.' These people are a lot worse off."*, p. 133), and feeling that one was not alone in facing certain health problems (Ian: *"You're not alone [...] it just makes you human"*, p. 145).

With respect to the comparative aspect, several interviewees admitted intentionally looking up others online with the explicit aim of feeling better about themselves, particularly in the context of weight and dieting:

> *If [...] you're always looking at pictures of models, then you might view yourself against what you view on the Internet. So I always read stories about obesity and think I'm doing pretty well.*
>
> – **Nicolas, 19**

> *I look up things and then realise how lucky I am. But that could be that I don't want to do any exercise, so I will look up these people that haven't done exercise or whatever, or look really awful. [...] I've always been 8.5 stone, and now I'm maybe 9.5, 10. So I look at people that are 13 stone and I think, oh, I'm really quite healthy.*
>
> – **Helen, 43**

Several interviewees also felt vindicated that their symptoms in fact were real. Michael and Miriam both felt that the Internet gave them a right to genuinely experience certain symptoms, which made it easier for them to accept their situations:

> *My heart every now and then just really hurts, and then it beats differently [...] When I was looking online, it says I've got irregular heartbeat. [...] So I am in a situation where I'm seeing what it is, because I am reading all the symptoms, but when I go to the doctors, [...] there's nothing there. So I don't know if it's a placebo effect that I'm having or if it's actually genuine. But I'm kind of in a situation where I just go for it online.*
>
> – **Michael, 22**

> *When I read about the amount of people [...] that get depression after [having a stroke] I was relieved in a way because I thought it was just me being silly – that there was something wrong.*
>
> – **Miriam, 59**

It is worth adding that "comparing and feeling normal" is not limited to the online realm. Both users and non-users described offline situations that could have taken place online in a similar way:

> *[My friend told me she was going to be incontinent] so I just cracked up laughing and said "Oh, join the club." There's hundreds and thousands of us in the same situation, but why have these sort of things still got the shame-bearing feeling? [...] Half of the problem goes away, just by talking about it and turning it into a joke and swapping ideas with everybody how to cope with it. The doctors don't tell you, the nurses don't tell you, you find out for yourself. But you have to talk about these things openly.*
>
> – **Charlotte, 78**

However, particularly for embarrassing conditions such as incontinence, there may be no open offline discussion due to the higher barriers against raising the topic. Moreover, the Internet is not limited to individuals' social circle, and through being extended to a wider network may include more people with certain conditions, particularly for conditions without a higher age-related prevalence.

Despite most interviewees describing feeling "normal" based on reading upon the condition online, some interviewees also thought that reading and hearing constantly

about a condition or treatment online also prevented feelings of normality. For example, Martha explained that she did not want to allow her IVF treatment to occupy too much space in her life:

> *I had a look [at fertility support groups], but I didn't participate, because [...] I didn't want to get too obsessed with the treatment and stuff. So I didn't really get too involved online as such.*
>
> – **Martha, 34**

Similarly, for Isabelle, belonging to a Facebook support group constantly reminded her of her son's autism, although living with him had already became normal for her. She did not want to be constantly reminded about it through notifications:

> *I just kept getting notifications, and I ended up just turning notifications off. Because I just felt, I didn't need to keep reading it. It's almost like, instead of reading it, I rather go to it when needed, instead of seeing what other people posted in the [autistic spectrum disorder] club.*
>
> – **Isabelle, 38**

Isabelle's example shows how "normality" may be a subjective experience, with changed information and support needs dependent on the condition and the stage the individual is at, as will be discussed in subsection 6.1.4.

5.2.3 Exchanging best practices with others

While the previous aspects affected individuals' health perception, the Internet also enabled them to tap into user-generated health tips that changed health behaviours. This particularly applied in the context of weight and dieting, where "what works" may depend on the individual. Both learners and sceptics were very interested in exchanging informal health advice online:

> *I used to go to these pages [...] for those who want to lose weight. So we kind of share our opinion about how to lose weight, healthy, without starving. What to eat, what not to eat – what works for me, what works for other people. [...] It's very helpful. I lost 20 kilograms. [...] And it makes me feel good when they say, "Oh, you lost so much weight."*
>
> – **Ahmad, 23**

In several cases, the Internet provided tangible advice on how to improve certain situations. Miriam learned on a forum that some simple changes to her footwear could ease her arthritis. At the same time, her quote also shows the close link to health service use:

> *Sometimes I would look it up on the Internet first, and it would say "do this and that, and that will relieve it," and it works. So I don't need to go the doctor and complain about something that I thought was really bad. So with my feet, I just need to wear different shoes and that's it. And so it does affect the amount of times I go to the doctors as well.*
>
> – **Miriam, 59**

The learners and sceptics in particular did not perceive lay advice as inferior to professional advice; this was also reflected in the motivations for health information-seeking in subsection 4.3.1. Danielle also confirms that doctors "see sickness only as disease and proffer [...] treatments that are technical 'fixes', whereas patients seek not only symptom relief, but also personally and socially meaningful explanations and psychosocial treatments for illness" (Kleinman, 1978, p. 78):

> *I could go to the doctor [...] but because they deal with people all the time, they don't understand – well, I guess they do understand – but they can't appreciate being in that situation at that time. Whereas if I get in contact with another mum who's going through this at this time, I would ask and comment on this group, and I'll see what they come back with.*
>
> – **Danielle, 30**

Of course, user-generated information may also be incorrect or create concerns, as in the case of Isabelle's father and the potentially fatal virus (p. 166). However, the more prevalent theme throughout the interviews was the realisation that particular health advice did not necessarily work for everyone:

> *I've been getting hot flushes, so I went on the Internet and looked up what's good for hot flushes and there's a few different things. I tried all of them – and none of them works! Obviously they work for different people, but they didn't work for me, so I'm just stuck with my hot flushes.*
>
> – **Miriam, 59**

In addition, providing user-generated advice on health networks created positive feelings of "mastering health questions", and giving something back to others. While Miriam established herself as an expert in her offline environment (p. 133), other learners such as Michael and Danielle acted as lay experts online:

> *I try to make the world a bit better, or try to open people's eyes, that's why I want to talk about problems, or healthcare problems. So that's what I mainly do online, just talking about things that people take for granted. [...] It's mainly with strangers, I don't ever have to face them.*
>
> – **Michael, 22**

> *I feel like I'm helping people because [they] will come on and say "I've got a four-month-old baby, they've broken out with eczema, I've been to doctor's, they've given me creams, it's not working," and straight away I would [...] say "Have you tried this, have you done that?" [...] Then they'd come back a few months later and go "Do you know what? It worked." [...] It makes you feel good for helping someone else going through that, so I think if you hadn't social media, I wouldn't be able to do that [...] because I don't know anyone with a child who has eczema and asthma.*
>
> **– Danielle, 30**

Danielle's last sentence also highlights the uniqueness of being an expert online. Particularly for less common health problems (and those that do not primarily occur with people of certain age groups, such as incontinence, p. 169), one may not necessarily know others affected by the specific condition.

5.2.4 Affecting perception of lifestyles

Several interviewees argued the Internet would affect their health-related lifestyle, as they looked up dieting (Brian, Hugh, Martha), keeping fit (Martha, Nicolas), or losing weight (Ahmad). As in the last section, interviewees reported that information and discussions online supported their healthy lifestyles:

> *I'm quite into sort of eating, eating clean as they say, so we do talk about that [in a Facebook group], I think that is healthcare as well. We swap our breakfast ideas, any little recipes that we find, and I follow a lot of [...] fitness people, and a lot of food people.*
>
> **– Martha, 34**

> *It's been beneficial in as far as [...] just normal dietary things – we're all quite ignorant sometimes and get things confused and mixed up, so for finding dietary properties of foods and things, it's just easy. Or to find an alternative to using cream, like some of the recipes for cooking.*
>
> **– Brian, 54**

While there was consensus among the interviewees about the positive health benefits of lifestyle-related information, the Internet also increased their sense of guilt about their lifestyles. As discussed in the context of the usage divide (subsection 4.2.3), the interviewees mentioned above also shared the rating that they gave their health – 4 or lower, based on being *"kind of averagely fit"* (Nicolas), *"sometimes [not doing] as much [as possible] fitness-wise"* (Martha), or thinking that *"it felt kind of prophetical to say 5, a little bit overconfident"* (Hugh). Hugh explained how the Internet has changed his awareness of fitness and health:

> *[The Internet influences health in] tangential ways such as pressure on fitness, health clubs, diet, the whole kind of lifestyle angle that the Internet is very filled with. [In the past] someone like me would never have bought a health magazine, you buy [a technical journal], and the Times or whatever. And now you can't avoid the kind of media awareness around lifestyle.*
>
> **– Hugh, 45**

Hugh also related to the tempo-spatial dimension in this context. This reflects how the Internet is "stretching [...] social systems across time-space, on the basis of mechanisms of social and system integration" (Giddens, 1984, p. 377). For Hugh, the Internet has decreased the boundaries between different areas of life, and thereby made lifestyle-related aspects around fitness and nutrition much more present:

> *You're always zooming to places that kind of promote certain values. [...] If you join a gym, you're immediately bombarded with emails from them about offers, courses [...] so their presence is felt not just when you're in the gym.*
>
> **– Hugh, 45**

While Hugh acknowledged that the increasing focus on lifestyle may be a general development of society, he argued that the Internet has contributed considerably to this development. He also admits experiencing guilt when not conforming to the active lifestyle propagated online:

> *I think the Internet has magnified that dialogue, it's amplified that dialogue. [...] It's this awareness, it's the faster media content that means that you're bombarded with stuff that leaves a passion in your mind about lifestyle, and so you have an awareness. You have a guilt if you're not having that lifestyle to some extent, in a way that maybe our fathers didn't.*
>
> **– Hugh, 45**

While this guilt did not affect Hugh's health service use, he confirmed that his health perception was influenced *"in several tangential ways"* through these lifestyle-modifying aspects of the Internet.

5.2.5 Managing identity and boundaries

Beyond Hugh's view of the Internet narrowing distance between places, the Internet also helped in the management of boundaries. As mentioned at the beginning of this chapter, for Ian the Internet is a place where he could *"join as a normal character without the mental health"* (p. 152), which gave him a space to feel free of the disease

and avoid *"being a target"* – his description of how he felt in the offline world. Joining online communities thereby helped Ian cope with his situation.

As another example, Kevin used technological boundaries to keep his disability unknown to his professional network. After his stroke a few years ago, he had to relearn how to speak and still clearly showed verbal limitations in the interview. Consequently, he used only messaging services to remain in touch with his industrial electrician colleagues, who – according to Kevin – did not know about his stroke. It was noteworthy that Kevin left his old voice as a voicemail. This may not have been an active choice, but shows how technology helped him to pretend to be healthy:

> *People I used to work for, I got a text from them [...] but I haven't told them I had a stroke. [...] But I won't text them I had a stroke, because like... I loved my job. No, I won't tell them. That cupboard is full of my tools, that's massive. [...] [But] I can't use them. I can't get back to work... apparently I'm disabled. [...] I was right-handed as well, so I can't even wire a plug up now. That's horrible.*
>
> **– Kevin, 40**

With his right hand cramped, writing on the phone required enormous effort, which he accepted as a necessary requirement. While Kevin did not describe the primary reason for keeping his disability to himself, his quote suggests that he may not want to admit his inability vis-à-vis his former colleagues, and expressed an underlying hope that he would be able to return to his job one day. In a different context, he said that the prospect of getting better one day kept him alive.

Another aspect of boundary setting is evidenced by Barbara, who was fighting her alcohol addiction. She said that viewing friends online was different from seeing them in person, because in person it was more common for them to have a drink – whereas at home, she would not be drinking on her own:

> *In a way I like to see them on Facebook, because if I see them at their place, we usually have a little social drink. [...] I thought I could cope, but last Tuesday I went out on a major drinking binge, so I'm back to square one now. Although I am not excessively drinking at home, the fact that I went over the top there, woke up and didn't realise what happened... it's made me realise I have got to give these [support] groups a chance.*
>
> **– Barbara, 41**

Barbara had therefore signed up to a local support group in her area. At a meeting with her four weeks later, Barbara said she had gone to the support group once, but was the only participant there and did not find it very useful. She was going to have another look at online support groups after the second interview.

5.2.6 Finding companionship

Apart from online support groups, the Internet helped several interviewees to find companionship. This related both to the technology itself and to the contacts made online, as captured by the argument put forward of Pols and Moser (2009, p. 160) that there is both a "love of the actual device" and an "affection because of its functionality". Several interviewees explained that the Internet helped them to cope with certain conditions and situations, and affected their outlook on life, both of which shaped their overall health perception. For example, the Internet helped Nathan to stay in touch with his children and grandchildren (see subsection 5.2.7), and to feel part of today's world through finding information and news online:

> *We are all on our own otherwise, so we find it a very good companion, very useful companion. [...] It fills a void, it fills a very large void to use this. [...] And it keeps us in touch with the outside world.*
>
> – **Nathan, 90**

While this was true for Nathan as one of the oldest interviewees, this also applied to Michael as the youngest. For Michael, the Internet helped him to get through some mentally challenging situations when he was younger:

> *When I was at a young age, I was affected by a lot. And that's how I got through it, it was actually through talking to people online. [...] I've been on websites, like where you can just talk to random people [...] they just want to talk, I want to talk, then we go from there really. [...] I didn't know who they were, but I got good advice from people.*
>
> – **Michael, 22**

While Michael valued that he could just find random people to talk to online about things that happened in his life, he also highlighted the affordances of the Internet that enabled this exchange online:

> *I've noticed it is a lot easier to talk to someone online than it is to talk to them in person. You can just express yourself a lot more really through the computer than in person. [...] I mean it's more or less the confidence that you can share it. Online, you can take your time, and reply whenever you want.*
>
> – **Michael, 22**

Similarly, Ian reported how the Internet provided him with additional friends who would listen and provide advice round the clock, particularly at night as *"that's when all the crazy thoughts go through your head."* Beyond the temporal dimension, Ian's quote also reflects the spatial aspect of *"an extra friend"* online:

> *If things are chasing you off and you can't really get to an answer, because you're a family man and your friends have heard it before, unfortunately whoever you're online with, and your good friends, they get to hear it. [...] It's just like an extra friend without the need of getting rained on or spending patrol in your car – it's amazing.*
>
> **– Ian, 42**

However, meeting others online also created negative experiences in a few cases. This related to hearing about losses online, for example:

> *[One of my online friends died] – and it's crazy, I don't even know him, and it's just what he brought to me, what he brought to our home, what he brought to our family – we've had so many conversations, but I've never seen him. I've never had a cup of tea with him, or gone out fishing with him. That's what I find fascinating about the Internet, it can really pull your emotional strings.*
>
> **– Ian, 42**

Negative experiences also referred to "trolling" online, which the Internet is prone to due to its relative anonymity and lower barriers compared to face-to-face contact (Wright and Bell, 2003). While Michael was not affected himself, he reported examples of how anonymity online prompted some people to bully others. Similarly, Susan explained that she did not participate online due to her fear of getting bad responses, partly due to her low spelling skills:

> *You get one comment and [...] everybody's jumping at you. I never write on there, unless I'm wishing a happy birthday. And I write "Happy birthday". [...] That's about my level. I don't have a conversation on there. I sometimes inbox someone privately so no one else can see what I'm writing, because I'm not a very good speller.*
>
> **– Susan, 58**

As in the other examples of companionship provided by Nathan, Michael and Ian, these experiences may not necessarily affect health directly. However, they may impact health ratings more widely due to their effects on self-esteem, how individuals view themselves and their mental state, and how they cope with their conditions.

Interestingly, those who were part of health-related online support groups – such as Ahmad for weight loss, Danielle for eczema and Isabelle for autism – did not mention companionship among the benefits they gained, but rather highlighted the exchange of information with others (see subsection 5.2.3). Then again, these interviewees also had strong support from their existing social environment. Similarly,

Whitney expressed the view that online support groups may be mostly for those who lack offline support and need to compensate for it online:

> *I know there are people that go to support groups, and they find it very useful if they're lonely, or they don't have friends. They need someone, I guess. [...] I have friends and family that I can use as support, I don't actually need other people, strangers.*
>
> – **Whitney, 60**

At the same time, Whitney explains how the offline support group provided *"someone to moan to"* for her husband with MS:

> *There are people there that [my husband] can moan to, he can say "oh, I've got this, and I've got that," and they actually understand, because the majority of people looking at him have no idea. [...] He looks fine, he doesn't look ill or whatever. But he is actually in pain all the time and can't walk, so people there understand.*
>
> – **Whitney, 60**

Several non-users also talked about the health benefits of emotional support from offline support groups. For example, Natalie explains how a group with other parents whose children suffer from hearing loss was helpful for her:

> *Obviously we're going through the same thing as they're going through, and it's nice to obviously share information with them and obviously if we can help one another, that's a good thing. [...] Obviously we're there for one another to get through it.*
>
> – **Natalie, 31**

However, Kevin and Elizabeth both argued that their offline support groups were less helpful due to the large age difference between the members, which highlights the importance of having commonalities between members of the support group:

> *I was young for having [my stroke], so it's always within the elderly section, and it wasn't always adequate. [...] They weren't useful, to me anyway. Because I was the baby there, being there in my 50s, the others were in the 70s and 80s. [...] So most of what they were looking at was different from what I was interested in.*
>
> – **Elizabeth, 55**

Elizabeth and Kevin both heard about online support groups for the first time in the interview. After being recommended a specific website, Elizabeth said she might give this a try when going to the library the next time. Kevin, however, re-stated that he *"hated computers"* (p. 112), and as a non-user potentially also lacked basic digital skills to access the groups.

5.2.7 Being embedded with family and friends

In addition, the Internet also helped with keeping in touch with family and friends. As with the companionship theme, this affected individuals' general outlook on life and self-perception, particularly in the context of depression:

> *If I didn't have the Internet, like Facebook or WhatsApp or whatever, it would just be on the phone, wouldn't it, in which you don't have a monitor. [...] There are times when I'm low, but I keep in touch with them on the Internet, and it will pick me up. So it's good for health reasons, as I'd probably be on a higher dose of medication [for my depression].*
>
> **– Barbara, 41**

Particularly older interviewees, as reflected in Nathan's quote about the Internet being a *"very useful companion"* and filling *"a very large void"* (p. 175), highlighted the capacity of the Internet to decrease substantially feelings of loneliness:

> *You're not speaking, you sort of lose track of things, you get introverted a bit, because loneliness is a terrible thing, and that is very dangerous to health, people being lonely. [...] I'm sure technology can help. A lot of older people do surf the net, as they call it, silver surfers, there's hundreds who do it, it's just that we never got around to it – we never needed to.*
>
> **– Irene, 82**

While Irene had one of her daughters around, and could reach the other one and her grandchildren via Skype with the help of the nearby daughter, Mildred confirms that loneliness is a problem, also with respect to discussing health-related matters:

> *The thing is that nobody wants to listen to you. I said to my son that I wasn't feeling well, and he said "You are alright, mom, what's the matter with you?" So I just leave it. Certainly I wasn't that ill, you just want somebody to talk to. That's what you want.*
>
> **– Mildred, 91**

Miriam also used the Internet as an outlet for easing her concerns without placing a burden on her children, thereby making her conversations with them more carefree:

> *It stops me from bothering my children as well, because they were frightened when it happened, and I don't want to keep saying, "oh, I had this and that"... so I would look it up on there, and it does explain everything.*
>
> **– Miriam, 59**

Then again, for communication with others nearby, the adigitals particularly highlighted that the quality of communication has deteriorated as a result of the rise of the Internet, thereby increasing loneliness:

> *Since people have done Facebook, people don't talk to people anymore. You don't pick up the phone like you used to and ring your friends, you just write a message on there, and I write you one back. Computers have stopped people from talking to each other and communicating.*
>
> – **Susan, 58**

This is also reflected in Sarah's self-critical assessment about her addictive use of the Internet in the past. After quitting Facebook, she says she has now established better social contacts, and gets more fresh air:

> *When I was addicted to Facebook I would never go out, I would go to work, come home, [...] straight on the iPad, tap it away. I didn't have a social life. [...] I think that's the problem with society today: most people spend most time hidden up indoors, instead of getting out there and getting a life. [...] Now, I spend more time out there than in here, [...] which makes me more healthy.*
>
> – **Sarah, 31**

Because the Internet can cause addiction and damage the quality of communication, it may also negatively impact health, as explored quantitatively in section 5.3.

5.2.8 Changing life circumstances

Finally, the Internet may also affect general life circumstances in times of health problems, or create opportunities which in turn indirectly influence health. Matthew, for example, had a spinal fracture in a work accident which caused him to be off work for more than a year. He says the Internet helped him to find out about benefits, overcome the hard times, and find a way back to work afterwards:

> *I am self-employed, and obviously, you got to feed the kids and pay the bills while you're off. [...] The most significant thing is finding out [about] benefits that you might get if you're lucky. [...] The employment agency [is] good at telling you [...] what you're maybe entitled to, but they're not as good as when you [...] find your own information on the Internet. [...] The difference between what they offer you and what you get if you bother to look is enough to make a difference to whether you manage.*
>
> – **Matthew, 47**

As in his previous example of how the Internet helped him to find words to use in consultations (p. 152), Ian also highlights the importance of the Internet for job interviews or writing a letter of complaint to his NHS consultant:

> *Knowledge, it's all this knowledge. [...] I didn't achieve too much at school, because mental health wasn't really understood in the 70s, you*

> *were classed as the class clown. [...] But the Internet has got me reading better, understanding words, nice words, nice lavish long big words that you can put into composition which help in job interviews, consultants, writing that letter of complaint [...] I'll be asking Google "Please can you show me the correct way how to draft out a letter," where to write and how to write it, and how to sign it off. So once again: Thank you, Internet.*
>
> – **Ian, 42**

Similarly, the Internet has fundamentally shaped Michael's interest in health. This has encouraged him to pursue a medical career, showing the close links between everyday life information-seeking (Savolainen, 1995) and job-related Internet use. By learning about healthcare online, he has decided to pursue a health-related National Vocational Qualification (NVQ), having previously worked in a design-related job:

> *Without the Internet, I wouldn't know half the things I do. [...] You can use it and educate yourself a lot on it. [...] But if I didn't read up on the topics, then I would not be doing [the new career path] at all.*
>
> – **Michael, 22**

While these instances may have affected health perception only indirectly, they show how the Internet can shape life circumstances that may influence health: the Internet helps people to overcome illness and enables social mobility. The importance of Internet use for social mobility has already been shown in other longitudinal research (Eynon et al., 2016, forthcoming) – and moving up in the social hierarchy may affect health both in terms of one's relative and absolute position, even in developed societies when controlling for other factors associated with inequalities (Bottero, 2004).

5.3 Quantitative models of selected mechanisms

Based on this qualitative exploration of how Internet use may relate to perceived health and health service use, some of the mechanisms can be modelled based on the available secondary data. The quantitative data by no means captures the richness of the qualitative data about the complicated ways in which these influences occur, but serves to follow up some of the identified relationships on a larger scale – while the qualitative data also helps to prevent SEMs leading to "inappropriate conclusions [when] statistical criteria are permitted to drive analysis and override substantive understanding of a problem" (Chin and Todd, 1995, p. 237).

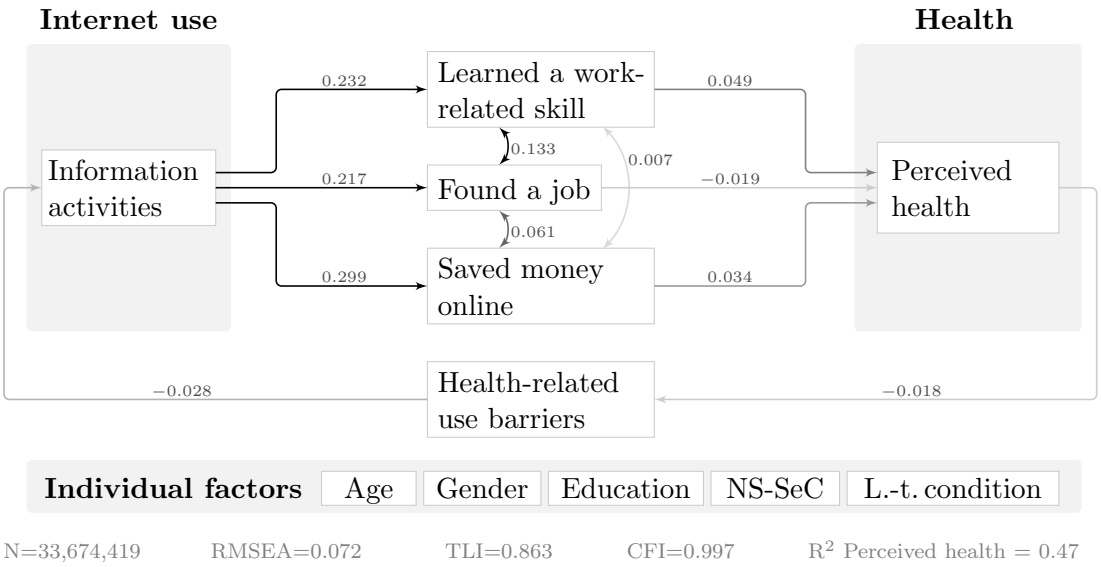
While the modelled paths follow the identified mechanisms, they also integrate influences observed in related mechanisms where appropriate. To this end, the

models are organised along the major themes of general (subsection 5.3.1), social (subsection 5.3.2), and informational (subsection 5.3.3) use-related outcomes. Although the previous sections discussed some parallels to mechanisms that also non-users experienced, this section focuses only on users. There are substantive reasons for this, as RQ_2 is primarily about users, but it is also a practical requirement as most of the necessary variables in OxIS (2013) are captured for users only.

5.3.1 Health outcomes of general usage

Figure 17 picks up the final mechanism presented in the previous subsection (5.2.8): changing general life circumstances. The available constructs largely cover those aspects addressed by the interviewees: learning a work-related skill (Q.O13a in OxIS, 2013) in relation to gaining subject-specific knowledge (Michael, p. 180) or learning words to use in job interviews (Ian, p. 179); finding a job (Q.C22a in OxIS, 2013) as an extension to Ian's point mentioned above; and saving money online (Q.C22b in OxIS, 2013), although the construct rather relates to consumer products online than learning about benefits, as in Matthew's example (p. 179).

FIGURE 17: SEM for "affecting life circumstances" mechanism



NOTE: Some paths have been omitted to improve clarity: all paths from the 5 individual factors *(age, gender, education, NS-SeC, long-term condition)* to all variables *(information activities, saved money online, found a job, learned a work-related skill, perceived health and health-related use barriers)*. Odds ratios and covariances for *health-related usage barriers* are not provided throughout this section to improve visual clarity.

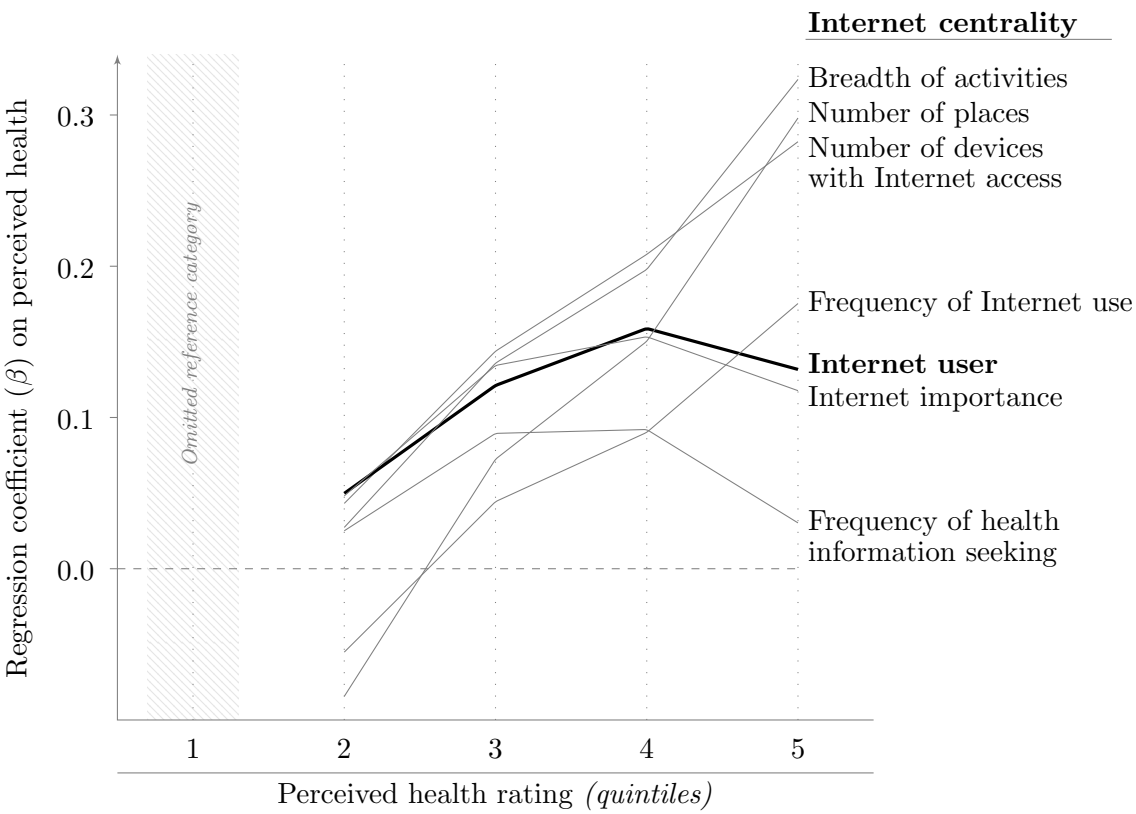
The model shows that information activities are positively related to these three outcomes of Internet use ($\beta = 0.232, 0.217, 0.229$), while some in turn are related to higher levels of perceived health ($\beta = 0.049, 0.034$, controlling for age, gender, education, socioeconomic status and long-term health conditions). Only finding a job was related to lower perceived health. While this may sound counter-intuitive, finding a new job may be related to unemployment or difficulties in the old one. Interactions with NS-SeC and unemployment were explored, but adding these effects considerably worsened the model's goodness-of-fit. No further variables or interaction effects were added, as the qualitative data did not indicate any particular influences that would mediate or moderate this relationship.

The three concepts of Internet effects show positive covariances, indicating that these outcomes are associated and may mutually reinforce themselves. With two of the interviewees who brought up this theme being learners (Michael, Ian), running the model in Figure 17 for the learners only shows that particularly the effects for "learning a work-related skill" ($\beta = 0.052$) and "saving money online" ($\beta = 0.042$) on perceived health become slightly stronger, which indicates that the learners may benefit especially from these Internet-based effects.

While all effects are moderate in size, the interviewees used phrases such as *"the most significant thing"* or *"the difference to whether you manage"* to illustrate the Internet's significance to changes in their lives. This is line with Eynon et al. (2016, forthcoming), who showed that the Internet can influence social mobility and general life circumstances. These shifts in social mobility can influence how individuals perceive their health (Bottero, 2004), which is corroborated by the interviewees bringing up these stories when asked how the Internet affected their health.

A different aspect around general usage is the qualitative finding that some interviewees (Hugh, Martha, Nicolas and Ahmad) rated their health status as lower based on not doing enough in terms of fitness and nutrition, with Hugh explaining that the Internet creates *"a guilt if you're not having that lifestyle"* (p. 173). This links in with the finding of Whitacre and Brooks (2014) that broadband availability was negatively associated with people describing themselves as in very good health, but positively associated with "good health". Figure 18 shows seven regressions for constructs around the centrality of the Internet in everyday life. In contrast to all other analyses in this research, the outcome variable of perceived health is split up into quintiles (1–3.36, 3.56, 3.70, 3.86–5), so that the Internet use concepts could be regressed on each level of perceived health through a multinomial logistic regression (controlling for sociodemographic covariates).

FIGURE 18: Multinomial regression for "affecting lifestyle" mechanism



NOTE: Each Internet centrality concept was included in a separate multinomial logistic regression on perceived health (split into quintiles) as the outcome variable, controlling for age, gender, education, NS-SeC, long-term condition.

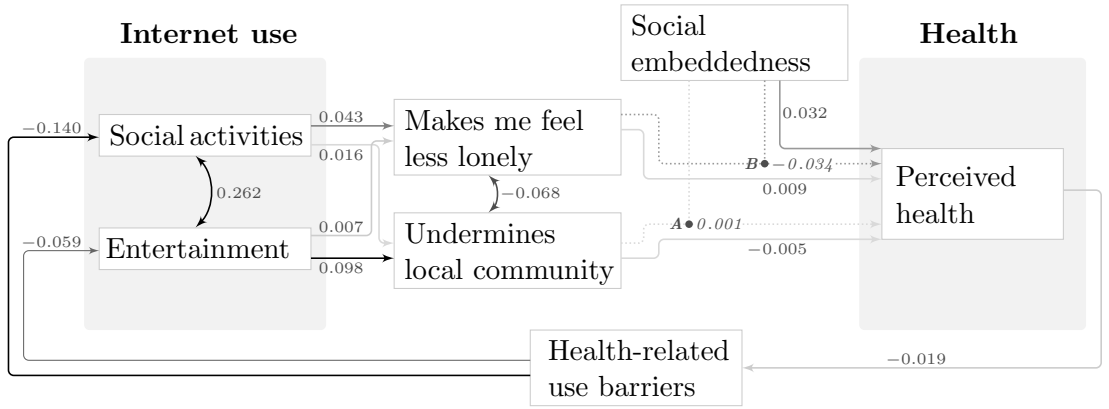
Figure 18 shows a monotonically positive relationship between Internet use and perceived health. However, being a user, placing a higher importance on the Internet, and frequent health information-seeking were linked less strongly to a rating of 5 ("very good"). In particular, the finding about use versus non-use (marked in bold) supports the idea that Internet use and perceived health are not always positively related, as the access divide suggested (see section 4.2). However, causality may be bidirectional: those with poorer health may search for health information more frequently (Houston and Allison, 2002), and the Internet may be particularly important for information and entertainment for those with poorer health (Rubin, 2009), even though long-term conditions are controlled for and healthy individuals may also search for health information (Cotten and Gupta, 2004). Moreover, Blaxter (1990) argues that younger people often assume a fitness perspective of health. The qualitative data supports this, as Brian aged 54 was the oldest interviewee expressing this view, whereas the quantitative data monotonically becomes more negative with

increasing age (rating 2: $\beta = -0.52$; 3: $\beta = -0.79$; 4: $\beta = -1.01$; 5: $\beta = -1.41$). Of course, the quintiles only roughly replicate the 1–5 scale.

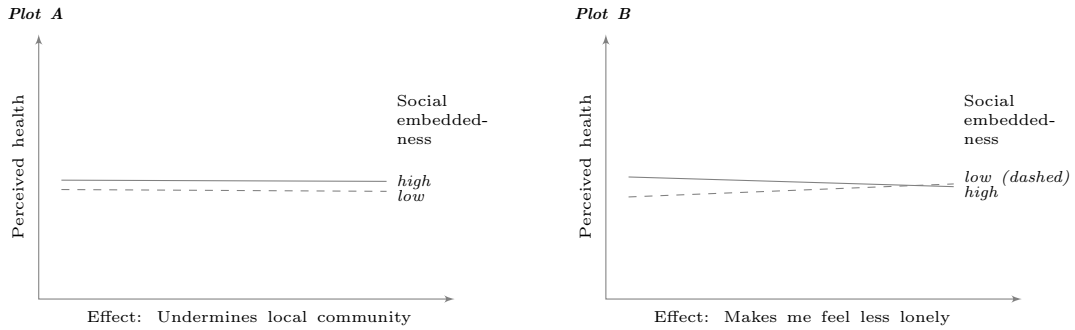
5.3.2 Health outcomes of social connections

The next set of mechanisms relates to individuals' social environment. Figure 19 includes the effects of "When I'm online I don't feel lonely" (Q.I8h in OxIS, 2013), with the Internet *"filling a void"* (Nathan, p. 175) or providing people with connections to others to talk to (Michael, p. 175). The effect of "the Internet undermining local communities by isolating people" (Q.B3a in OxIS, 2013) extends Susan's point that *"people don't talk to people anymore"* (p. 179). Both paths may be moderated by social embeddedness based on the qualitative findings and previous research (McKenna and Bargh, 1998; Bessière et al., 2008).

FIGURE 19: SEM for "finding companionship" mechanism



Interaction effects



NOTE: Some paths, covariances and odds ratios are omitted (see Figure 17).

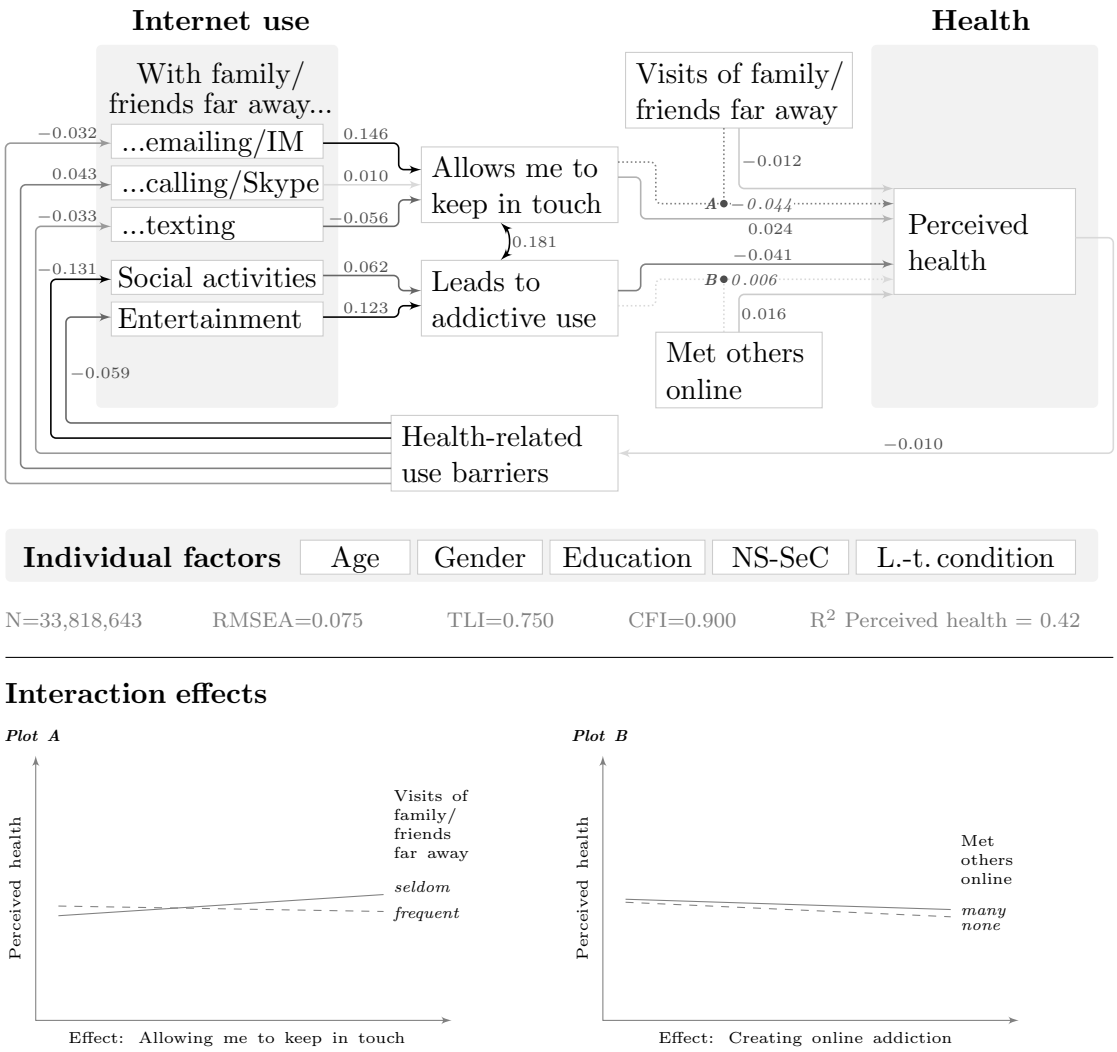
Indeed, although people also build up relationships to their devices as such (Pols and Moser, 2009), social activities in particular are associated with decreased loneliness ($\beta = 0.043$). Entertainment activities are primarily related to viewing the Internet as undermining local communities ($\beta = 0.098$). Both outcomes are related only weakly to perceived health ($\beta = 0.009$ and -0.005). However, the relationship between perceived health and agreeing that the Internet reduces loneliness depends on the level of social embeddedness ($\beta_B = -0.034$): in line with McKenna and Bargh (1998) and Bessière et al. (2008), the Internet as a way of reducing loneliness had particularly positive benefits for the more lonely (plot B). Social embeddedness itself was associated with higher perceived health ($\beta = 0.032$), as detailed in section 6.2. For the weak negative association between the Internet undermining local communities and perceived health ($\beta = -0.005$), social embeddedness played no role ($\beta_A = 0.001$).

Figure 20 provides an SEM related to being embedded with family and friends (see subsection 5.2.7), evaluating both positive outcomes through "Going online allows me to keep in touch with people" (Q.I8a in OxIS, 2013), based on Barbara's (p. 178) account, and "I feel I can be addicted to the Internet" (Q.I2e in OxIS, 2013), based on Sarah's Internet addiction, which decreased her real-world contacts with family/friends (p. 179). These two effects are modelled as the results of activities around being in contact with family/friends far away and social/entertainment activities respectively. The SEM introduces an interaction effect with face-to-face visits of family/friends far away, which moderates the relationship between being in touch with friends/family and positive health outcomes – as Susan indicated, this depended on actually knowing someone far away. Another interaction effect of meeting new people online (Q.C11a–j and Q.C12 OxIS, 2013) moderates the relationship between addictive use and perceived health; as Michael (p. 175) explained, the Internet helped him cope with health-related and general life events (despite leading to some addictive use).

The results show that, after controlling for sociodemographic covariates, being in touch with family/friends is associated with higher perceived health ($\beta = 0.024$). This relationship is stronger for those who only seldom have face-to-face contact with their family/friends, while it is nearly non-existent for those with frequent face-to-face contact ($\beta_A = -0.044$, plot A). Online addiction is negatively associated with perceived health ($\beta = -0.041$). While those who met someone online have slightly higher levels of perceived health, the relationship between addiction and perceived health is barely mediated by meeting others online ($\beta_B = 0.006$, plot B). In contrast

to the qualitative results, keeping in touch via email has a stronger relationship than Skyping, although this may partially be due to the OxIS construct not clearly separating phones from video-enabled Internet services (see Appendix IV, p. 301). In line with Dobransky and Hargittai (2006), health-related barriers are negatively associated with social activities online (except for calling).

FIGURE 20: SEM for "being embedded with family/friends" mechanism

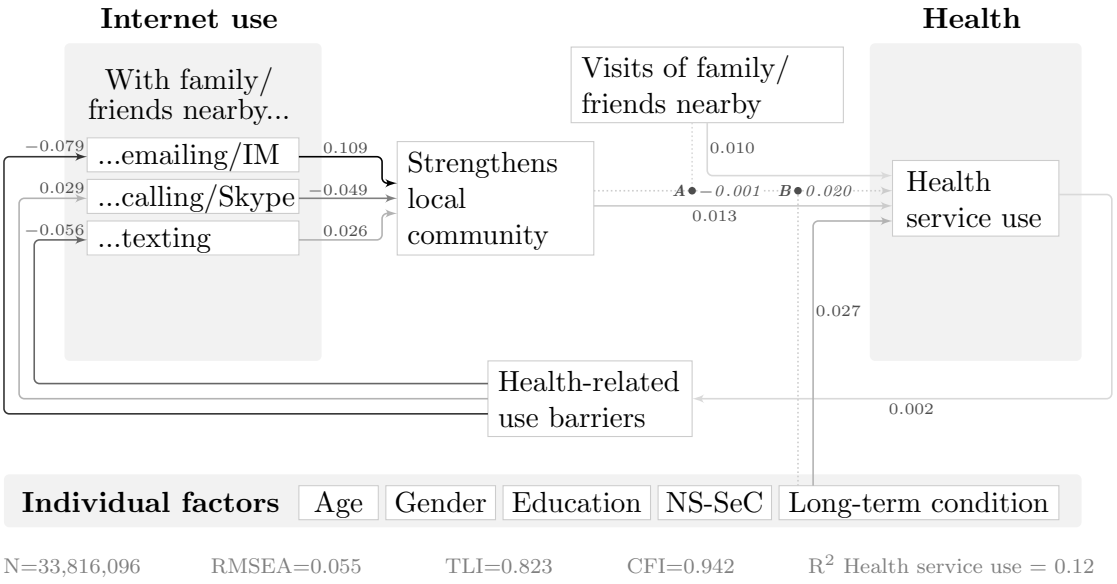


NOTE: Some paths, covariances and odds ratios are omitted (see Figure 17).

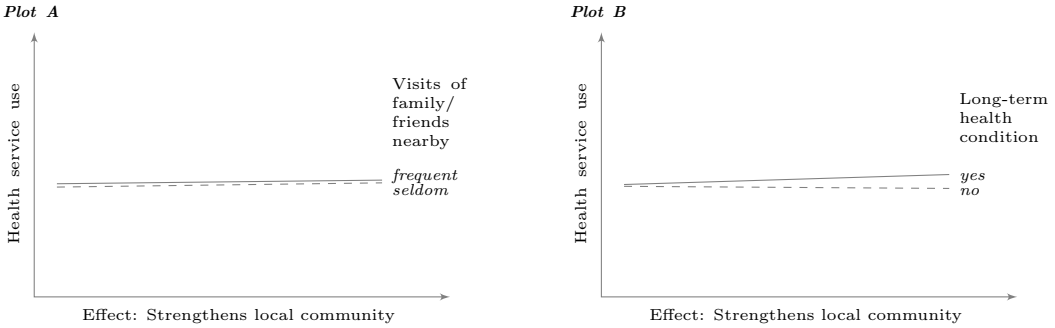
The last mechanism rebuilt with the quantitative data is monitoring for emergencies. Based on the qualitative data, being monitored by nearby family members or friends affected their health service use by allowing them to get help when needed.

This was expressed by Miriam (p. 159), Kevin (p. 159), Charlotte (p. 159), and – with a slightly different slant – Nathan (p. 159). Figure 21 shows the SEM including the ways of communication with individuals nearby (email, calling/Skype and texting) and their effects on the perceived strengthening of local communities. Face-to-face interactions with nearby family/friends and long-term conditions potentially moderate this relationship; particularly those who had experienced a stroke (Miriam, Kevin) or suffer from chronic, potentially life-threatening conditions such as heart failure (Nathan) brought up this theme.

FIGURE 21: SEM for "monitoring for emergencies" mechanism



Interaction effects



NOTE: Some paths, covariances and odds ratios are omitted (see Figure 17).

Particularly emailing/instant messaging ($\beta = 0.109$), and texting to a certain extent ($\beta = 0.026$), were related to saying that the Internet strengthened local

communities. Those who agreed with these positive community effects created by the Internet in turn also made more use of health services ($\beta = 0.013$). In line with the qualitative results, this effect was slightly stronger for those with a long-term health condition ($\beta_B = 0.020$, plot B). These effects were not moderated by face-to-face visits, indicating that either method of monitoring and caring (in person or mediated by technology) may have similar results.

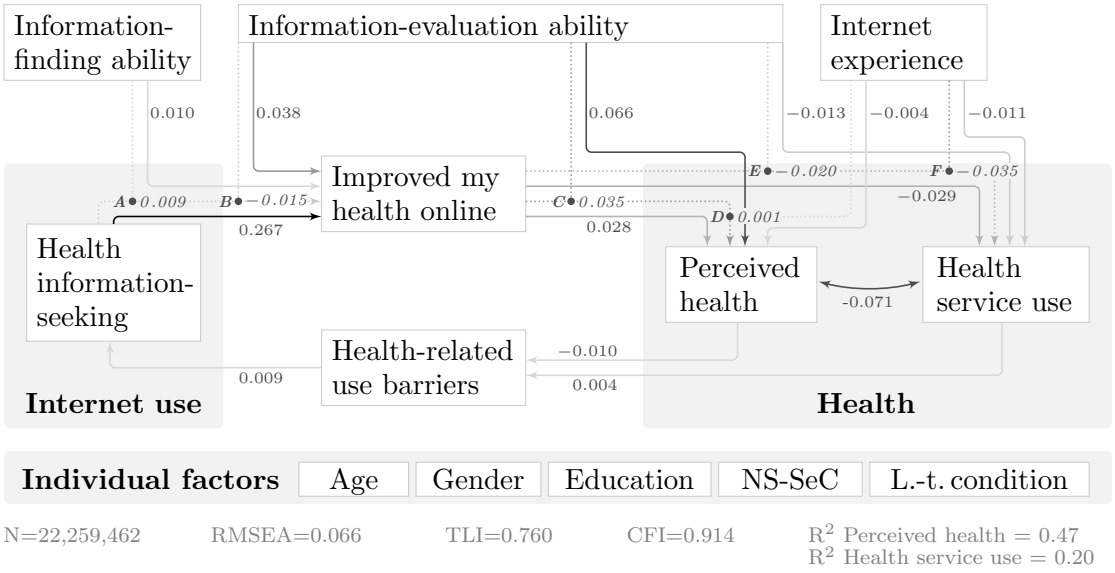
5.3.3 Health outcomes of health information-seeking

This final subsection returns to the effects of health-related Internet use, and therefore revisits and extends the model of health information-seeking introduced for the usage divide (Figure 11). In addition to the previous model from Chapter 4, Figure 22 incorporates some aspects that stood out as moderating factors in the qualitative data: information-finding skills (Q.I2g,h/Q.I8f in OxIS, 2013) that influenced whether individuals found relevant health information; and information-evaluation skills (Q.C23a in OxIS, 2013) necessary for searching and interpreting health information online (Martha, p. 120; Margaret, p. 154). Finally, Internet experience (Q.C3n in OxIS, 2013) in terms of the length of time that an individual had used the Internet (see Appendix IV, p. 297) is included in the model, as the Internet could *"exacerbate how you're feeling tenfold, 9 times out of 10 exacerbated the urge to see a GP"* (Ian, p. 157).

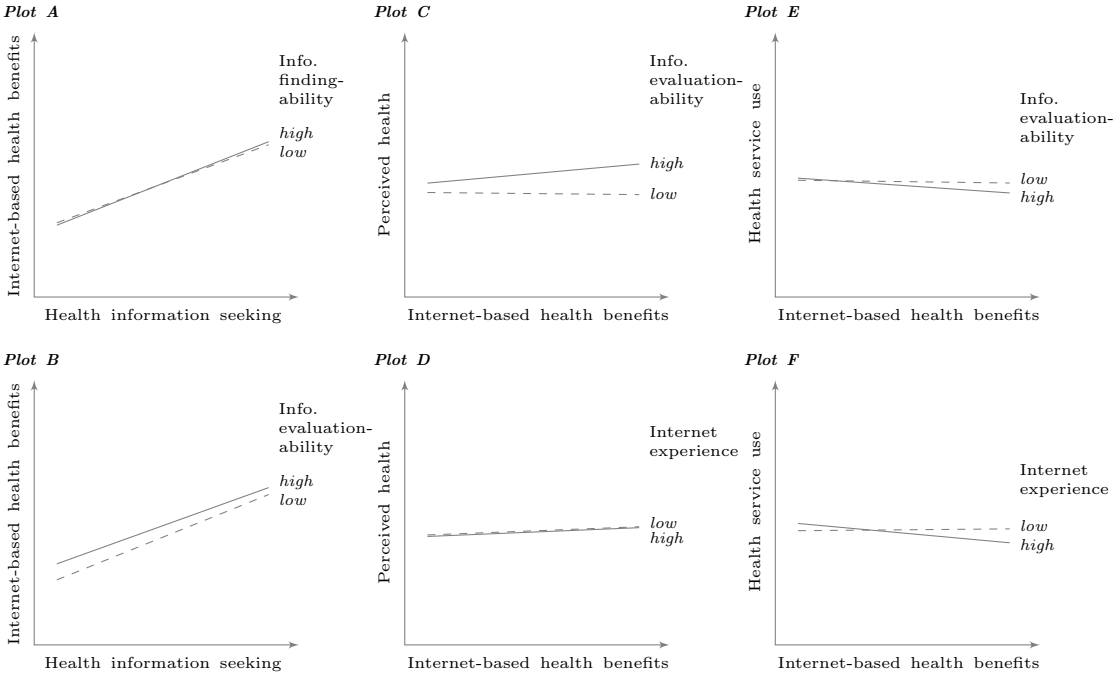
As discussed before, higher levels of health information-seeking are related to higher chances that individuals found something they considered beneficial for their health ($\beta = 0.267$). The SEM shows that those with higher information-finding and evaluation skills are more likely to say that they found something online that improved their health ($\beta = 0.010$ and $\beta = 0.038$), although the interaction effect with health information-seeking is very weak ($\beta = 0.009$ and $\beta = -0.015$, plots A/B). Moving to the effects of Internet use, finding something online that improved health was related to higher levels of perceived health ($\beta = 0.028$) and lower levels of health service use ($\beta = -0.029$). Supporting the qualitative findings, both relationships are stronger for those with higher information-evaluation skills ($\beta_C = 0.035$, plot C; $\beta_E = -0.020$, plot E). Internet experience only weakly contributes to both health outcomes ($\beta = -0.004$ and $\beta = -0.011$). However, in line with Haythornthwaite and Wellman (2002) and Ian's argument about experience offsetting some of the Internet's influences, Internet experience mediates the effects of the Internet on health service use ($\beta_F = -0.035$): as plot F shows, while for those with low experience, the relationship is basically inexistent, those with Internet

experience had lower levels of health service use through finding information online. Other ways of conceptualising Internet experience (for example, through breadth of activities online or frequency of activities online) were explored and yielded similar results.

FIGURE 22: SEM for health information-seeking mechanisms



Interaction effects



NOTE: Some paths, covariances and odds ratios are omitted (see Figure 17).

As presented throughout this chapter, some of these mechanisms may differ for mental health and stigmatised conditions, as in Ian's example about his men's health condition that he had kept secret for 30 years (p. 151), Miriam's example of her stroke-related depression (p. 169) and Charlotte's non-user example of incontinence (p. 169). Running the model shown in Figure 22, with health service use for mental health conditions only, shows that Internet use does not lead to a reduction, but rather a very weak increase ($\beta = 0.005$), whereas the relationship ceased to exist for stigmatised conditions ($\beta = 0.000$, both not shown in figure) such as mental health, sexually transmitted diseases or incontinence (see Appendix VI, p. 307). This indicates that particularly for these kind of conditions, the Internet may encourage health service use (Berger et al., 2005) – or, based on the rather weak relationships, at least not reduce health service use (as for the general concept shown in Figure 22).

These quantitative conceptualisations showed that many of the effects observed in the qualitative interviews also exist quantitatively on a larger scale, and confirmed some interesting interaction effects based on the qualitative findings and the previous literature. This was the case even though the quantitative data captures only a fraction of the complexity that became evident in the qualitative data presented in sections 5.1 and 5.2. The findings from both parts of the mixed methods design will be discussed in relation to the previous theoretical and empirical literature in Chapter 7, with section 7.1 relating the mechanisms identified back to the analytical framework set up in Chapter 2. However, before moving to the discussion, what is left for the final results chapter is a perspective on the individual factors that have been omitted in most of the SEM visualisations, as well as the social and geographic environment that has been touched upon on several occasions throughout this chapter.

Results 6

Individual and contextual influences

> *The way I always look at this is, when someone's born, it's a clean slate. It's like a canvas when you're drawing something, and then it's design: the way you treat it, and bring it up. That's how you're going to be in the world and [...] that's what actually makes a person who they are today.*
>
> – **Michael, 22**

In his own words, reflecting his background in graphic design, Michael explains how both the family and the local area profoundly shape who he is and how he acts. In this context, he referred particularly to his interest in knowledge, which is reflected in his classification as a learner, but also to the values he obtained from his parents, for example around valuing doctors, caring for people, and trusting others. Then again, Michael was bullied in his neighbourhood when he was younger, as in that area *"you just [got] harassed for walking down the street."* As a consequence, Michael argued that because *"people tend to struggle a lot more"* in the area that he used to live in, *"their health, they got a higher risk for all this than others."*

While the precise effects between sociodemographic, area-related and wider social characteristics may be hard to disentangle, the interviewees implicitly or explicitly expressed different thoughts about how Internet use may influence their health perception and health service use based on their backgrounds – or showed certain traits that influenced this relationship against the odds of their backgrounds. Some sociodemographic differences have already been reflected in the typology of health information-seekers in Chapter 4 (Table 13, p. 129), although they were not

predictive of these types (Table 14, p. 130), and in the mechanisms in Chapter 5 (Figures 17–21, p. 181–187). This chapter extends these insights by focusing on the mechanisms used throughout Chapters 4 and 5, questioning *how* these individual characteristics may shape motivations and enablers, and what role further attitudes, the social environment and area-based factors play, as RQ_3 seeks to understand.

Understanding these individual and contextual factors is important from various perspectives. From a theoretical perspective, it provides a basis for the discussion around agency and structure, with those factors presented in this chapter as structures that may enable and restrict agency (Giddens, 1984). From a practical perspective, sociodemographic and area-based characteristics are tangible starting points that policymakers and practitioners can build on, and link into other policy priorities such as social inclusion. From a methodological perspective, as most influences presented in this chapter were part of the SM process used to generate the quantitative data, it is appropriate to scrutinise the way in which they matter.

The main point argued and illustrated in this chapter is that individual and geographic factors primarily influence factors antecedent to Internet use and health outcomes separately – whereas the social environment additionally influences their relationship. To this end, this chapter primarily focuses on health information-seeking, building on Chapter 4, and draws in insights about other mechanisms from Chapter 5 where appropriate, with all regressions based on users only. The following sections discuss individual (section 6.1) and social environment factors (section 6.2). As the qualitative data created most insights for this part, and various constructs were available in the simulated data to test these, these two sections present the qualitative data first, and then follow up with a quantitative analysis of emerging relationships. In contrast, for the geographic environment, the qualitative data helped to explain some of the quantitative data, so section 6.3 presents the data in that order. This results chapter therefore takes advantage of the flexible possibilities that the bracketed mixed methods research design enables (Greene et al., 1989).

6.1 Demographics

Interviewees referred to age, education, socioeconomic status and long-term conditions when explaining their Internet use, health and health service use, or differences became apparent during data analysis. Some characteristics were more prevalent among certain sociodemographic groups, while some also made a difference to health *despite* sociodemographic features. However, this was not true for gender. While

primarily older women referred to skills as inhibitors of Internet use (Hargittai and Shafer, 2006), no outspoken or implicit gender effects were observed. Accepting that the initial assumption of a link between gender and health could not be sustained, gender has been left out as a separate subsection (Berg and Lune, 2014). It is still included as a control variable in quantitative analyses, as gender has an influence in some respects based on the quantitative data (see Appendix V, p. 302).

In addition, for all quantitative analyses in this chapter, the bidirectional relationship between Internet use and health discussed in Chapter 4 should be kept in mind, although the feedback loop's effects are reduced by controlling for long-term health conditions in all analyses. In contrast to previous chapters, this chapter focuses on regressions as opposed to SEMs, primarily looking at the effects of health information-seeking ("Did you ever find anything online that helped improve your health?", QC.22d in OxIS, 2013). This makes it possible to model the relationships with more complex interactions, while keeping the models clear and parsimonious, and helps to improve replicability based on other datasets.

6.1.1 Age

Beyond the well-researched effect of health decreasing with age, which is reflected in the data (see Appendix V, p. 302), this research largely confirmed Blank's (2015) "first law of the Internet" about younger people using the Internet more frequently than older people. Interviewees mentioned a range of interrelated age, period and cohort effects, partially already presented in the context of the "finding companionship" mechanism (subsection 5.2.6), which will be detailed in this subsection.

First of all, Internet interest declined with age, as reflected in George's statement about not being *"bothered by looking at the high velocity stuff"* anymore (p. 121). While many non-users attributed non-use to their age or generation at first, they often revealed a lack of interest afterwards. Irene, a delegator, for example, does not use the Internet mainly due to a lack of interest – but also due to the excellent possibilities available to her to outsource her questions, which reflects Reisdorf et al.'s (2012) point about a social network decreasing the desire to become an Internet user:

> *We're not too old, if we wanted to do it bad enough. You can do anything if you really want to. It's willpower. But if you're not interested, there's no point. [...] We don't want it badly enough. You can go down the library if you want to. And as I say, our daughter and the granddaughter, they show us everything that we need to know. So we don't need it.*
>
> **– Irene, 82**

While there may be both age effects in terms of ability and the need for technology, and period/cohort effects due to not gaining familiarity with the Internet at a younger age, Nathan as a 90-year-old user also explains how general interest is more important than age in response to why he uses the Internet, in contrast to others of his generation. He illustrates that the elderly do not necessarily form one homogeneous group that barely uses the Internet (van Deursen and Helsper, 2015):

> *I'm just interested, I've always been interested. I like innovation, I like new things.*
>
> **– Nathan, 90**

For users, age also made a difference in terms of their activities online (Dutton and Blank, 2013, van Deursen and van Dijk, 2014) across all user definitions, as discussed in subsection 4.2.3. Matthew, for example – referring to himself as an old person at the age of 47 – explains that both in general and in the context of health information, the prominence of the Internet as a first port of call may also be lower for older people:

> *Being an old person, you have to remind yourself. Youngsters now, first thing they do, if they need to know something, [they are] straight on [the Internet]. It's just their first point of call, whereas for older people like me, we actually grew up with dictionaries and things.*
>
> **– Matthew, 47**

At the same time, even Anthony, who described himself as relatively versatile online and found the Internet very important for general information, felt that his generation could justify seeing a doctor without finding out online first:

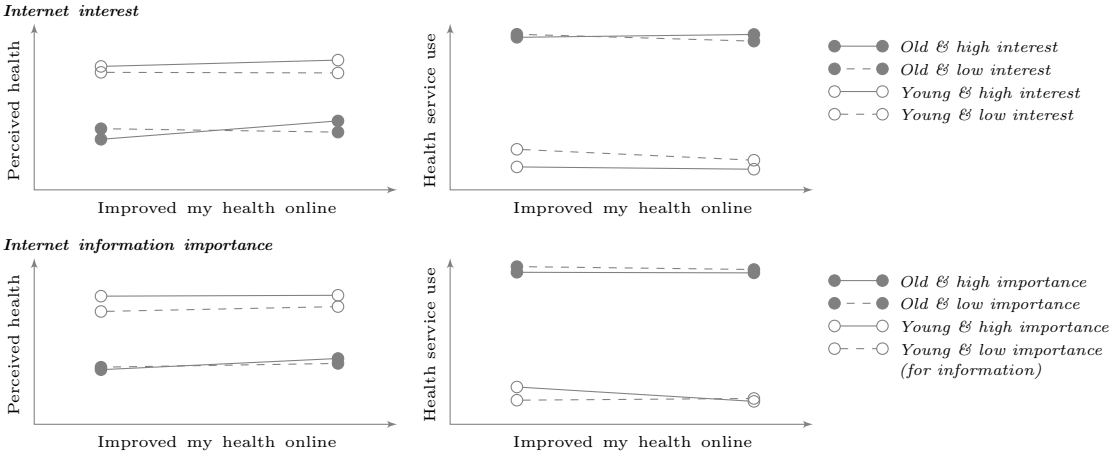
> *I'd go straight to the doctors. I'm of that age where you're not driven by the Internet. My daughter, she's straight on her phone [...] and is "oh no, I woke up this morning and had a headache, do you think I've got a tumour?" I'm not of that generation. If I wake up and have a problem that I think is serious enough that it isn't going within the next few hours [...], I go to the doctor's and take it from there.*
>
> **– Anthony, 58**

From these qualitative insights, the two aspects that can be followed up quantitatively are Internet interest (Q.A9 in OxIS, 2013) and the importance of the Internet for information relative to other media (Q.A2a–d in OxIS, 2013), although it would be useful to have the latter specifically for health given Anthony's and Matthew's accounts. Table 16 includes both as moderators on the relationship between Internet-based effects and perceived health/health service use respectively.

TABLE 16: Regression for age effects

	Perceived health			Health service use		
Independent variables	β	β	β	β	β	β
Improved my health online	0.019	0.013	0.018	−0.020	−0.018	−0.015
× Age			0.008			0.008
Internet interest		0.017	0.016		−0.019	−0.019
× Age			−0.015			0.025
× Improved health			0.025			0.016
× Improved health × Age			0.013			0.001
Internet info. importance		0.027	0.024		0.005	−0.002
× Age			−0.020			−0.014
× Improved health			0.003			−0.011
× Improved health × Age			0.010			0.016
Age	−0.207	−0.200	−0.206	0.420	0.417	0.412
Gender	0.099	0.102	0.102	0.200	0.200	0.200
Education	0.324	0.328	0.327	−0.031	−0.030	−0.030
NS-SeC	−0.103	−0.098	−0.101	0.038	0.037	0.034
Long-term health condition	−0.485	−0.475	−0.476	0.032	0.033	0.034
Adjusted R^2	**0.48**	**0.48**	**0.48**	**0.21**	**0.21**	**0.21**
N *(in million)*	**33.90**	**33.87**	**33.87**	**33.90**	**33.87**	**33.87**
Largest CI	**2**	**2**	**2**	**2**	**2**	**2**

Interaction effects



The results show that Internet interest is related to higher perceived health ($\beta = 0.016$) and lower health service use ($\beta = -0.019$); particularly for older users, finding something that improved health online was related to higher levels of perceived health for those who were interested in the Internet (see upper left interaction plot). Similarly, the levels of perceived health were higher for those who

considered the Internet an important information source ($\beta = 0.024$), whereas the influence on heath service use was essentially non-existent ($\beta = -0.002$). However, particularly for younger users, considering the Internet as an important information source was related to lower levels of health service use if they found something that they considered improved their health (see lower right interaction plot).

In contrast to Blank's (2015) "first law of the Internet", with rising age and growing health problems, interest in health-related questions – and hence health information-seeking – may increase. However, lifestage plays a role, too: particularly after the birth of her children, Martha started taking her own health more seriously, and therefore increasingly turned to the Internet for health advice:

> *Before I had the girls, I sort of felt very indestructible, I didn't really worry about my health that much. Since I've had [children], I worry about obviously their health and also mine, so I have been looking up things.*
>
> – **Martha, 34**

Further supporting the life stage aspect, Hugh points out that despite increased interest in health topics with age and having children, the effects may be hard to separate due to the rise of the Internet over the last two decades:

> *Widespread Internet use is about the same length as I've had children. My first child was born in 2000, so widespread Internet use has come in what... 1998? [...] I've never really known the need to have medical advice without the Internet really. A man under 30 never goes to the doctor.*
>
> – **Hugh, 45**

The quantitative data confirms that the proportion of health information-seekers as a share of users goes up across life stages (students: 41%; (un)employed without child: 69%; (un)employed with child: 69%; retired: 72%), but intensity measured on a 1 to 5 scale is relatively stable at $\mu = 1.6$, with the exception of students, who look up health information most frequently ($\mu = 1.9$). Analysing the relationship between improving one's health online and perceived health across life stages does not yield an interpretable pattern. To a certain extent, this is reasonable – both Martha and Hugh, for example, referred to frequency rather than to effect or outcomes.

6.1.2 Education

Education and socioeconomic status are related to each other and, for some of the factors presented here, difficult to separate analytically. Nevertheless, there are two education-related factors of relevance to the Internet and healthcare: a generally

higher appreciation of the medical profession based on one's own professional experiences, and education as a way of finding and evaluating information online.

Regarding the first point, particularly those with higher education expressed the view that *"[doctors] know what they are doing – they're medically trained"* (Danielle, p. 148). Similarly, Michael voices trust in his physician:

> *I always trust them, because that's what they've studied [...] and built a life around a career, it's a very dedicated job to be honest. So I do take their word for it.*
>
> **– Michael, education level 4**

While Hugh attributed this high respect to his own professional experience (p. 134), Anthony highlighted how going through education yourself makes you aware of the difficulty of the training that medical professionals have gone through:

> *You got to trust them. [...] I don't really question that. Particularly with the little bit of education that I've gone through, [I] realise what they have done to become that. [...] I think the problem is that we expect people to be completely faultless. It just doesn't happen.*
>
> **– Anthony, education level 4**

At the same time, several interviewees with lower educational backgrounds expressed an attitude that a medical professional should "just know" without the need to look something up. Just as Barbara explained how doctors *"sometimes don't know themselves and they have to look into their big fat book"* (p. 118), Mary also thought that doctors should not have the need to look things up:

> *Doctors always check online, don't they? [...] They're on the Internet as well, so they're looking up things. And I'm like "you're a doctor, you should know this, you shouldn't need to look it up."*
>
> **– Mary, education level 1**

However, there are several stray instances from the equation "high education = high trust" which can partially be traced back to the typology of health information-seekers. Ethan (education level 4), a sceptic, thought that medical professionals lacked deep knowledge, could not keep up to date, and were not interested in him as a patient (p. 149). Then again, Kathleen (education level 0) showed very high appreciation of her doctor and valued how he googled things for her (p. 141); this is characteristic of the delegator type. It has to be acknowledged, however, that formal education levels may not always be the best indicator of actual education for the elderly (see subsection 3.2.2).

Moreover, in line with the general link between education and health literacy (Parker et al., 2003), education helped individuals to evaluate information online, for example in terms of knowing about the mechanisms behind Wikipedia, or in terms of corroborating information from several sources:

> *[On Wikipedia] sometimes you [think] "well, that's not right," but [...] someone will go and will say "that's not right," and they will correct it or flag it up that it's wrong, so it's kind of self-regulating. But you have to take it a little bit with a pinch of salt [...] and you have to know that.*
>
> – **Matthew, education level 3**

Although Anthony rarely looks up health information because he would worry too much, he explains why education helps to evaluate the quality of information:

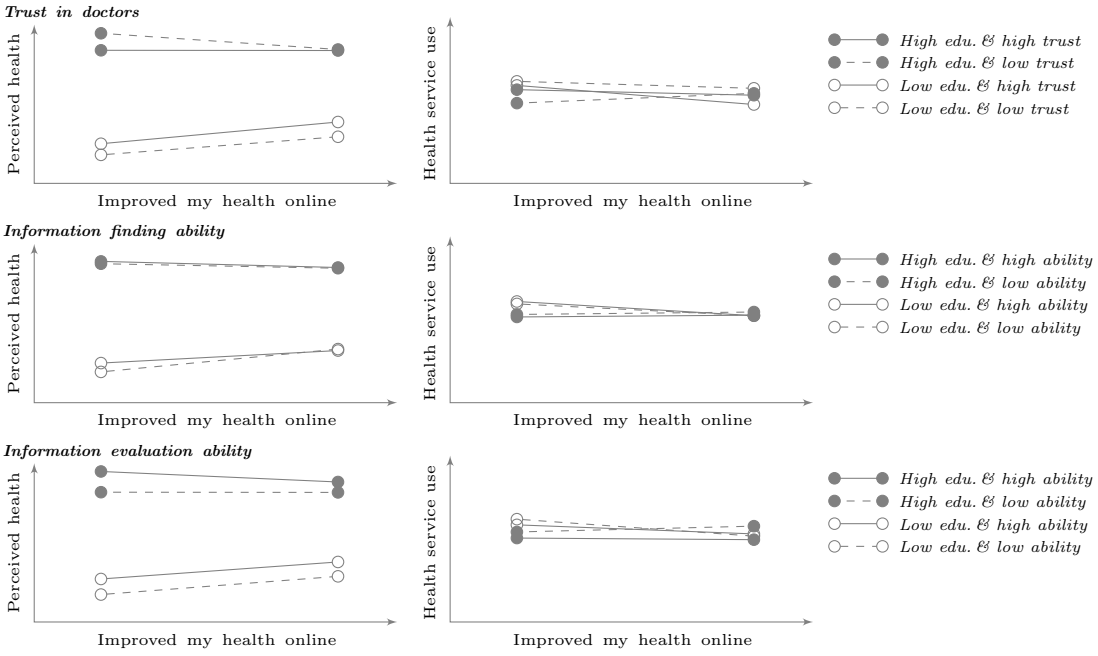
> *For the rest of your life, you never look at things quite the same. [My dad] would always say "Oh, have you seen that in the paper?" [...] He doesn't even consider that if he'd look at another paper, that same story would tell you something completely different. [...] You can't do that once you've been to university. You got to look at three or four and then make up your own assessment. [...] If you don't do a degree, you'll never get like that. That inquiry in mind doesn't happen naturally.*
>
> – **Anthony, education level 4**

These qualitative insights are modelled quantitatively in Table 17, breaking down the relationship by analysing interactions with trust in doctors (Q.A7b in OxIS, 2013), information-finding ability (Q.I2g,h/QI8f in OxIS, 2013) and information-evaluation ability (Q.C23a in OxIS, 2013). Both trust in doctors and information-finding ability are barely related to either outcome concept ($|\beta| < 0.015$), whereas information-evaluation ability is related to higher perceived health ($\beta = 0.051$) and lower health service use ($\beta = -0.019$). The three-way interactions show that for those with lower levels of education, having higher trust in doctors and higher information-evaluation skills is related to higher levels of perceived health. For information-finding skills, there are barely any differences, which – in line with the qualitative data – indicates that information-finding largely corresponds to differences in education. Effects on health service use are less clearly visible, but the interaction plots confirm that education in combination with information-finding and evaluation skills, as well as trust in doctors, is related to lower levels of health service use. In general, the differences in the magnitude of education between both outcomes ($\beta = 0.309$ and -0.025) are a corollary of HES data not allowing differentiation between education levels (only by age/gender).

Table 17: Regression for education effects

	Perceived health			Health service use		
Independent variables	β	β	β	β	β	β
Improved my health online	0.019	0.016	0.020	−0.020	−0.018	−0.019
× Education			−0.049			0.026
Trust in doctors		0.013	0.007		−0.002	−0.007
× Education			−0.035			−0.000
× Improved health			0.018			−0.023
× Improved health × Education			0.009			−0.003
Information-finding ability		0.015	0.010		−0.004	−0.003
× Education			−0.003			−0.006
× Improved health			−0.011			0.001
× Improved health × Education			0.007			0.006
Information-evaluation ability		0.053	0.051		−0.020	−0.019
× Education			0.000			−0.013
× Improved health			−0.017			0.000
× Improved health × Education			−0.008			−0.005
Age	−0.207	−0.198	−0.201	0.421	0.417	0.418
Gender	0.099	0.107	0.100	0.200	0.197	0.202
Education	0.325	0.318	0.309	−0.030	−0.028	−0.025
NS-SeC	−0.103	−0.096	−0.101	0.038	0.035	0.035
Long-term health condition	−0.485	−0.487	−0.489	0.032	0.033	0.033
Adjusted R^2	**0.48**	**0.48**	**0.49**	**0.21**	**0.21**	**0.21**
N *(in million)*	**33.90**	**33.90**	**33.90**	**33.90**	**33.90**	**33.90**
Largest CI	**2**	**2**	**2**	**2**	**2**	**2**

Interaction effects



Finally, individuals also used online resources to compensate for a lack of education and to appear more sophisticated in consultations (see subsection 5.2.8). Ian, who says about himself that he was an underachiever at school because *"mental health wasn't really understood in the 70s"*, described how he has used the Internet to mask his low education (p. 152): to learn keywords to use during a consultation, and to look up keywords afterwards. In addition, he also realised that if he wanted to be heard, he could not *"send off a shirty letter"* to his NHS consultant without guidance from the Internet and its spell-check features.

6.1.3 Socioeconomic status

Differences in socioeconomic status mainly related to technology attitudes, but also different feelings surrounding health responsibility and views about what can be done to treat certain conditions. For general Internet use, socioeconomic status was related to a lack of financial resources to get a device or pay for a connection.

In terms of technology attitude, several interviewees of high socioeconomic status, such as Hugh (p. 135), demonstrated strong views that costs in the health system may need to be reduced where possible without undermining quality, whereas those of lower socioeconomic status had strong views against using technology in healthcare. There is a substantial overlap with the adigital type, as reflected in the lower socioeconomic status of this group (Table 13, p. 129). Sarah, for example, argues that due to the lack of personal contact, she might go to hospital straight away, where the technology that her GP used was not implemented yet:

> *If you had a serious problem [...], you can't actually speak to someone. Your next step is just to go straight up to the hospital and speak to someone there. But then they'll probably don't have a receptionist, just some touch screen board to book yourself in. [...] And if they do it [at the GP's], they're going to do it to the hospitals as well, to save money up there as well. They're always looking for ways of cutting money. [...] You don't get the personal touch anymore – you're just a bloody number.*
>
> **– Sarah, NS-SeC 3**

Sarah also put forward the argument that doctors and nurses get paid, but that services on the telephone would not count:

> *[The doctor] gets paid to do a job, he gets paid to come and see you, or you to see him, you should be able to talk to him. [...] [They have] all these touch screen book-in screens, and rather than [the receptionist] do her job, what she's being paid for, [she says] "go and use that."*
>
> **– Sarah, NS-SeC 3**

Those with a lower socioeconomic status also showed lower understanding for why they would need to look up things, exhibiting a lower level of ownership of their health. While this may partially relate to education, the majority of interviewees to whom this applied were of a lower socioeconomic status:

> *I'd rather go to the doctor straight away. [...] Instead of reading that it's going to be bad on the Internet, you can just go to the doctor's, and the doctor can explain to you properly. [...] So best to go to the doctor's first, or hospital if you're that worried.*
>
> – **Mary, NS-SeC 3**

However, this does not necessarily mean that these individuals used health services more frequently. In addition, Sarah's quote again provides evidence for the low self-efficacy in terms of finding out about an issue online instead:

> *If I feel I need to go [to the doctor's], then I will. I usually wait a bit longer and if it gets worse then I go. It's like going to the dentist, you will wait, until it gets really really bad you just don't go there. [...] No, I haven't yet googled it, if the problem continues then I might.*
>
> – **Sarah, NS-SeC 3**

The interviewees also demonstrated varying health beliefs. Those with higher NS-SeC talked about the Internet helping them to learn about preparing healthy meals (Brian and Martha, NS-SeC 1) or exercising (Martha and Hugh, NS-SeC 1). The beliefs about what could be done about conditions also differed fundamentally. Danielle attributed high value to online support groups for buffering against depression (although she did not use any during her pregnancy as her sister is a midwife):

> *I can quite easily see, if [...] I was a single mum, and I was down in the dumps, that support group would have definitely helped me. Because it feels like, they're going through what you're going through, and they can empathise with you [...] and it's better than anything at all, so I can see how it would help depression and things like that.*
>
> – **Danielle, NS-SeC 1**

In contrast, Mary argued that drugs were the only solution for treating mental conditions such as ADHD and depression:

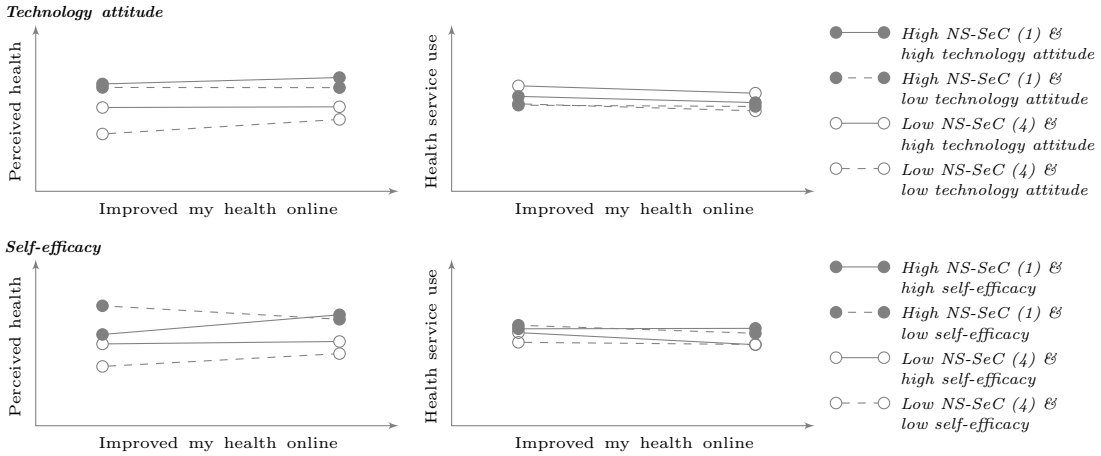
> *ADHD, you have to be diagnosed with it, you have to be treated for it, but there's not a lot else they can do, they can only like give you medication and hope that you deal with it. [...] Depression, you just get tablets for it, there's nothing else you can do.*
>
> – **Mary, NS-SeC 3**

Table 18 shows these concepts analysed quantitatively. Technology attitude is particularly positively associated with perceived health ($\beta = 0.044$). Self-efficacy has a rather weak main effect ($\beta = 0.009$), but shows an interesting interaction effect: for those who found something online that improved their health, self-efficacy was related to higher perceived health for those of any socioeconomic status. For health service use, both main effects and interaction effects are rather weak.

TABLE 18: Regression for socio-economic status effects

	Perceived health			**Health service use**		
Independent variables	β	β	β	β	β	β
Improved my health online	0.019	0.017	0.019	−0.020	−0.018	−0.019
× NS-SeC			0.008			0.006
Technology attitude		0.048	0.044		−0.014	−0.014
× NS-SeC			0.021			−0.019
× Improved health			−0.006			−0.004
× Improved health × NS-SeC			−0.018			−0.004
Self-efficacy		0.012	0.009		0.007	0.009
× NS-SeC			0.049			−0.007
× Improved health			0.019			−0.001
× Improved health × NS-SeC			−0.037			0.016
Age	−0.207	−0.192	−0.193	0.421	0.417	0.416
Gender	0.100	0.105	0.101	0.200	0.199	0.200
Education	0.324	0.316	0.307	−0.030	−0.029	−0.027
NS-SeC	−0.103	−0.099	−0.108	0.038	0.037	0.040
Long-term health condition	−0.486	−0.485	−0.485	0.032	0.032	0.031
Adjusted R^2	**0.48**	**0.48**	**0.49**	**0.21**	**0.21**	**0.21**
N	**33.90**	**33.90**	**33.90**	**33.90**	**33.90**	**33.90**
Largest CI	**2**	**2**	**2**	**2**	**2**	**2**

Interaction effects



Finally, in addition to the effects of socioeconomic status around the wider outcomes of Internet use (mentioned above), for example in terms of finding a job, learning a work-related skill or saving money (see section 5.3), lower socioeconomic status may also be a reason for not having access to the Internet. Natalie, for example, said she did not use the Internet partially because it was too expensive (p. 121), while Courtney found it difficult to find the money to pay for her mobile data package. Health problems may also cause socioeconomic status to go down, making the Internet less affordable, as described by Elizabeth (p. 109) and followed up from the perspective of long-term health conditions in the next subsection.

6.1.4 Long-term health conditions

As introduced at the beginning, long-term health conditions were a continuing source of digital exclusion (Dutton and Blank, 2013), with only 42% of those with a long-term condition using the Internet, compared to 79% of the general population (OxIS, 2013). Similarly, the group of adigitals, which largely consisted of non-users, had the highest proportion of those with long-term conditions, both as a share of the total population (47% compared to $\leq$ 23% for other groups, Table 13) and as a share of the users (26% compared to $\leq$ 18% for other groups).

For those with long-term health conditions who used the Internet, those affected by conditions that caused concern for the future (such as a stroke) or considerable discomfort (such as eczema) and who had high self-efficacy found the Internet a good resource to find out more about their condition and to become more involved. This became apparent in Miriam's reason for looking up health information (*"just to see what you can do to ease it a bit, or if there's anything that can help you"*, p. 132), Elizabeth's desire to learn more to challenge the doctor (p. 163), or – for other family members' long-term conditions – Danielle's efforts to find out about additional treatments online to tackle her son's eczema (p. 163), and Whitney's investigation of state-of-the-art treatments for MS (p. 118). As discussed throughout Chapter 5, the information acquired online enabled these four interviewees with high self-efficacy in terms of self-care and interpreting symptoms (self-efficacy: 4 on a 1–5 scale) to evaluate urgency (Miriam, p. 171 and p. 158), and shape discussions in the medical consultation (Whitney, p. 161; Elizabeth, p. 163; Danielle, p. 163).

It is worth highlighting that this was true across types, even if there were differences between types: whereas learners, worriers and sceptics with a long-term condition tended to be more explorative in their Internet use, pragmatists were predominantly instrumental (such as Brenda looking for schools for her autistic son

rather than reading about autism, or Matthew looking up monetary benefits rather than reading about the condition which made him unfit for work for 2 years). At the same time, worriers had rather low self-efficacy by the definition of the group, but still showed different approaches to managing their long-term conditions: Helen, with lower self-efficacy, denied her epilepsy to herself (p. 138), while Natalie, with higher levels of self-efficacy, rather actively looked for ways to cope with her partner's diabetes (p. 161).

Then again, the importance of self-efficacy largely related to conditions that the individuals considered severe and worth doing research on as they could do something about them. Carol explains that she did not look up anything about her thyroid after she obtained the right medication dose, just as Matthew did not look up much information about his blood pressure problems, while both also show the relatively instrumental approach typical of pragmatists:

> *I don't think you need to [look up more about hypothyroidism], really. Because it can be controlled very much by drugs, you just have to take it for the rest of your life, that's all [...], once they get the medication right, that's it, you can carry on.*
>
> **– Carol, hypothyroidism**

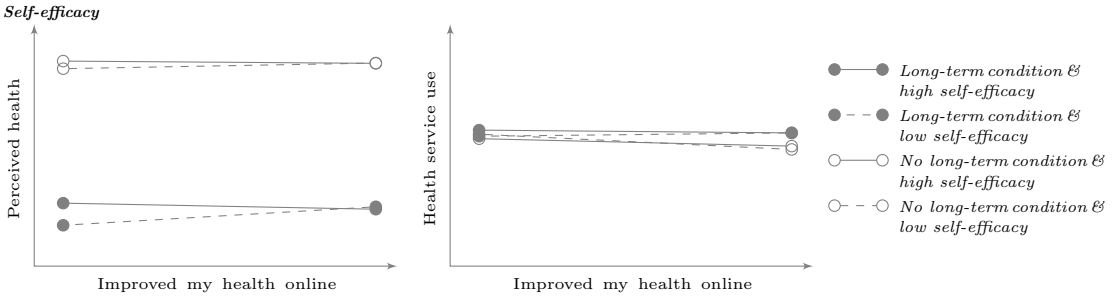
> *I think if it's untreated, high blood pressure builds up to a big problem, but blood pressure isn't actually like a big disease, you can just live with it, it's no big deal really.*
>
> **– Matthew, spinal fracture and high blood pressure**

This particular aspect of a self-efficacy is followed up in Table 19. A more suitable analysis would include more details on the kind of long-term condition involved, and would potentially evaluate not only the individuals' own long-term conditions, but also those of family members or close friends. The results show that once individuals found something that improved their health, their levels of perceived health were nearly the same independently of their self-efficacy (with pronounced differences in the overall level of perceived health between those with a long-term condition and those without). Similarly, for health service use, there is hardly any difference in the level of self-efficacy. To a certain extent, this is in contrast to the qualitative data. However, as already discussed in subsection 4.3.3, the self-efficacy concept in OxIS captures a general belief about whether life in general can be influenced, rather than any specific beliefs about getting a certain action done, as Bandura (1986, 2004) had in mind.

TABLE 19: Regression for long-term health condition effects

	Perceived health			**Health service use**		
Independent variables	β	β	β	β	β	β
Improved my health online	0.019	0.020	0.020	−0.020	−0.019	−0.020
× Long-term health condition			0.008			0.019
Self-efficacy		0.020	0.021		0.005	0.004
× Long-term health condition			0.010			0.006
× Improved health			−0.028			0.002
× Improved health × L.-t. con.			−0.014			−0.012
Age	−0.207	−0.207	−0.210	0.421	0.421	0.421
Gender	0.099	0.099	0.100	0.200	0.201	0.202
Education	0.325	0.323	0.321	−0.030	−0.031	−0.031
NS-SeC	−0.103	−0.103	−0.106	0.038	0.038	0.038
Long-term health condition	−0.485	−0.485	0.487	0.032	0.032	0.031
Adjusted R^2	**0.48**	**0.48**	**0.48**	**0.21**	**0.21**	**0.21**
N *(in million)*	**33.90**	**33.90**	**33.90**	**33.90**	**33.90**	**33.90**
Largest CI	**2**	**2**	**2**	**2**	**2**	**2**

Interaction effects



It is worth highlighting that those with very good medical support (Nathan, p. 123) or knowledge within the family, particularly for hereditary conditions, also did not necessarily use the Internet to seek out more information about their conditions, and neither did they experience any outcomes:

> *My mum and dad had [high blood pressure], and obviously you're like "Mum, so what did you get?" and all that, you kind of talk about it first, but kind of it's just normal now.*
>
> – **Matthew, spinal fracture and high blood pressure**

> *I didn't look that up [...] All the members of the family have had glaucoma, so I did know a little bit about that.*
>
> – **Karen, glaucoma**

Moreover, those with a long-term condition and a high learning attitude also established themselves as experts in their condition (Ziebland, 2004), as was characteristic of the learner type more generally. Just as Danielle provided expert advice on eczema online (p. 172), Miriam constantly offered those in her offline social environment to *"give them [her] point of view once [she has] read up on it on the Internet"* (p. 133) for her own and other people's conditions.

These aspects concerning receiving family members' advice and providing information to others show the interrelated nature of individual factors and the social environment. Further ways in which the social environment contributes to the relationship between Internet and health will be discussed below.

6.2 Social environment

Apart from the Internet context, many interviewees thought of their social environment (family/friends, including carers) as a source of support and information. This related to their general wellbeing, their information needs for hereditary conditions, lay knowledge about domestic remedies, and tangible support in everyday life, reflecting previous findings on patient experiences (Hughes et al., 2013).

As the previous chapters showed, the social environment was important in the context of the Internet in several ways: Chapter 4 made it clear that supported users had higher perceived health and lower health service use than non-supported users, just as those with proxy access or availability rated their health higher than the fully excluded (Table 12 and Figure 10). The finding that even proxy availability was related to some outcomes suggests that the social environment may contribute to health. Chapter 5 confirmed that the Internet, as a technology enabling users to connect with friends and family, may have positive effects on health, while the Internet was also a source of finding companionship, particularly for those with low social embeddedness (see Figure 19).

What has remained largely untouched so far is the role of the social environment in health information-seeking, particularly in terms of transforming online activities into outcomes. In contrast to the individual factors discussed in section 6.1, the social environment had two major functions for the relationship between Internet use, perceived health and health service use: compensating for a lack of enablers, which facilitated using online resources and achieving outcomes (Helsper, 2012; Eynon and Malmberg, 2012), and filtering information or buffering its potentially negative effects. Both of these are discussed in the next two subsections.

6.2.1 Compensating for a lack of enablers

The interviewees provided a range of instances where children, grandchildren or neighbours helped them to use the Internet by showing them how to do certain basic activities online, just as Shirley would *"ask a million questions and get help"* from her neighbours and children (p. 111).

In addition, those who had no device presented several ways of getting health information from the Internet (as in the case of Kathleen, p. 151), using the Internet at somebody else's home with the help of somebody else (Natalie, p. 121), or getting a *"proper device"* instead of just their own mobile phone with Internet access (as in the case of Wanda, p. 121). For Natalie, as someone who did not possess her own device to access the Internet, her offline network also offered a way of exchanging information from offline and online sources:

> *We've got this diabetic book, [if] she hasn't seen it, then we say "Oh, we've got this book, do you want to use it?" and then she'll say, "Oh, I've got this off the Internet, do you want to try this?" and obviously we share different things [...] to manage the condition.*
>
> **– Natalie, 31, non-user**

Internet users also specifically received help with finding (health-related) information online. Ethan, for example, explains how his friendship with medical doctors helped him to find the right words to put into a search engine, thereby allowing him to muddle his way through towards finding an answer:

> *If you can't find what you want because you're not using the right terminology on Safari, then to begin by talking to somebody who was trained as a [medical doctor] many years ago is the next best step. Then you find out that they question their own knowledge, and say, "Well, why don't you try and look it up on the Internet?" [...] It's another key word, but you then... you know, successive approximation.*
>
> **– Ethan, 69, user**

In contrast to Ethan, a sceptic, other users with medical knowledge did use the expertise of a medically trained family member or friend right away, such as David, whose wife worked in a pharmacy, or Courtney, whose husband was a paramedic. While particularly for embarrassing questions that one may value that *"no one can see what [one is] doing"* (p. 152), this was not the case for those interviewees who had medically trained family members.

In addition to finding information, the social environment also helped in interpreting information and turning advice provided online into action. For example,

Sarah explains how her mother told her to go to the GP as quickly as possible (although Sarah would often ask her mother before actually searching online):

> *I [always] ask my mother. She's pretty much "Miss Know-it-all" when it comes to certain things – she pretty much knows it all. [The last time I asked her] she turned around and said "Get your arse down to the doctors", so I went down there and I made an appointment.*
>
> – **Sarah, 31, user**

Similarly, Mary prefers to delegate health information-seeking altogether due to a lack of information-evaluation ability, as she would always *"take it as the worst thing ever"* and would then be *"sitting there panicking"* (p. 147). By asking her family, she would get actionable advice instead of a pile of unrelated facts:

> *[My family] would look it up, and then get back to me, and let me know what it sounds like, and then I'd say I'd go to the doctor's if it was anything bad. They say always something like this and I'll be like "Okay, I'll put myself in to the doctor's."*
>
> – **Mary, 29, user**

Interestingly, Mary also uses the information obtained from her family to be able to refer to them during the consultation, thereby allowing her to say *"my family says"* rather than "I looked it up online," which according to some interviewees (Ian, p. 160; Wanda, p. 161) and previous research (Murray et al., 2003; Lewis, 2006) was not well-regarded by physicians:

> *I leave that to the doctor to decide and tell me, but I do relate some of it back to the doctor, like, "My family's read this and says it's this," and the doctor's just like, "No, that's fine, your child is fine" or "You're fine."*
>
> – **Mary, 29, user**

At the same time, this responsibility shift also may have allowed her family to withhold some of the information that could have caused concerns for Mary. While Mary did not necessarily know about this (and did not mention it in the interview), other interviewees provided multiple instances where they filtered online information for their family members and friends, as will be presented in the next subsection.

6.2.2 Filtering information

Sometimes "less may be more", so that several interviewees deliberately filtered information for others to avoid them worrying unnecessarily:

> *My mom was in such shock that I was the sort of support mechanism. [...] I was the one that went to the computer, and if there was anything really*

> *negative and really harmful, then I was able to siphon that out, because it would have been too much for her. So actually I did do that.*
>
> – **Helen, 43, user**

This "social filtering" also happened in the case of Isabelle (p. 166), who found out that a virus might have fatal consequences for her dad, who had MS. In most cases, the primary motivation for withholding this information from those who were concerned was to avoid diminishing their positive spirit. In addition, Margaret recalls the situation of a friend with pancreatic cancer, where filtering plays a crucial role due to the low chances of finding hope-inspiring information online (Chapple et al., 2012). Margaret also did not share the information due to feeling awkward:

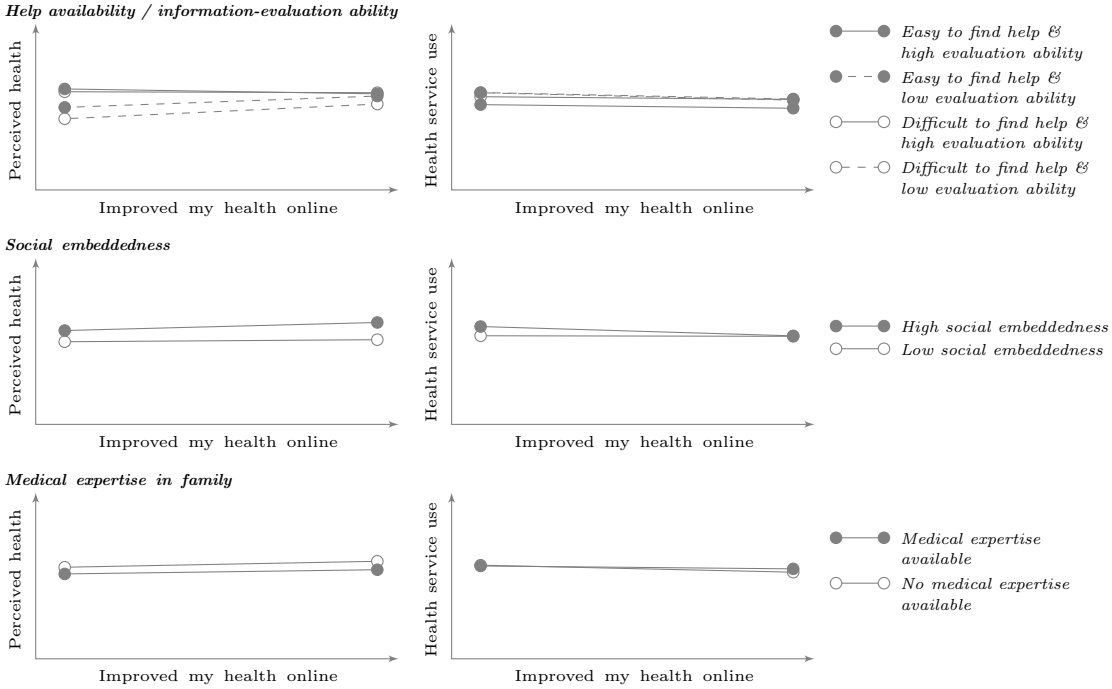
> *A friend of mine was diagnosed with pancreatic cancer, and [...] a few of us who were friends with her sort of looked it up, so that we could understand what she was going through in that. [...] [But] we just sort of talked amongst friends. It's an awkward one, something like that to talk to with somebody, because for a long time, she was convinced that she would get better – but she didn't.*
>
> – **Margaret, 57, user**

Table 20 models the effects discussed in this section through a three-way interaction with help availability and information-evaluation ability, and two-way interactions for social embeddedness and medical expertise in the household. For perceived health, help availability ($\beta = 0.028$), information-evaluation ability ($\beta = 0.050$) and social embeddedness ($\beta = 0.047$) are positively related to perceived health, whereas medical expertise in the household is instead associated with lower health perceptions. The interactions show that for those with low evaluation skills, once they found something online that they considered improved their health, the levels of perceived health are relatively similar – with slightly higher levels for those who found help more easily, which highlights the role of help and reconfirms the role of social embeddedness. Interestingly, for health service use, while help availability ($\beta = -0.013$) and information-evaluation ability ($\beta = -0.020$) are related to lower levels of health service use, social embeddedness slightly increases health service use ($\beta = 0.013$), perhaps as individuals are encouraged to seek help (see subsection 5.3.2). There are barely any interaction effects for health service use, although the upper right plot shows that those who found help more easily and are able to evaluate information well have the lowest levels of health service use.

Table 20: Regression for social environment effects

	Perceived health			Health service use		
Independent variables	β	β	β	β	β	β
Improved my health online	0.019	0.016	0.017	−0.020	−0.017	−0.017
Help availability		0.028	0.028		−0.013	−0.013
× Improved health			−0.006			−0.000
Information-evaluation ability		0.050	0.049		−0.020	−0.020
× Improved health			−0.027			0.007
× Help availability			−0.014			−0.014
× Help avail. × Improved health			0.000			−0.002
Social embeddedness		0.045	0.047		0.013	0.016
× Improved health			0.010			−0.015
Medical expertise in household		−0.027	−0.025		0.007	0.004
× Improved health			−0.003			0.008
Age	−0.207	−0.197	−0.199	0.421	0.417	0.419
Gender	0.099	0.107	0.107	0.200	0.198	0.199
Education	0.325	0.328	0.325	−0.030	−0.027	−0.026
NS-SeC	−0.103	−0.097	−0.099	0.038	0.036	0.036
Long-term health condition	−0.485	−0.472	−0.471	0.032	0.035	0.034
Adjusted R^2	**0.48**	**0.48**	**0.48**	**0.21**	**0.21**	**0.21**
N	**33.90**	**33.90**	**33.90**	**33.90**	**33.90**	**33.90**
Largest CI	**2**	**2**	**2**	**2**	**2**	**2**

Interaction effects

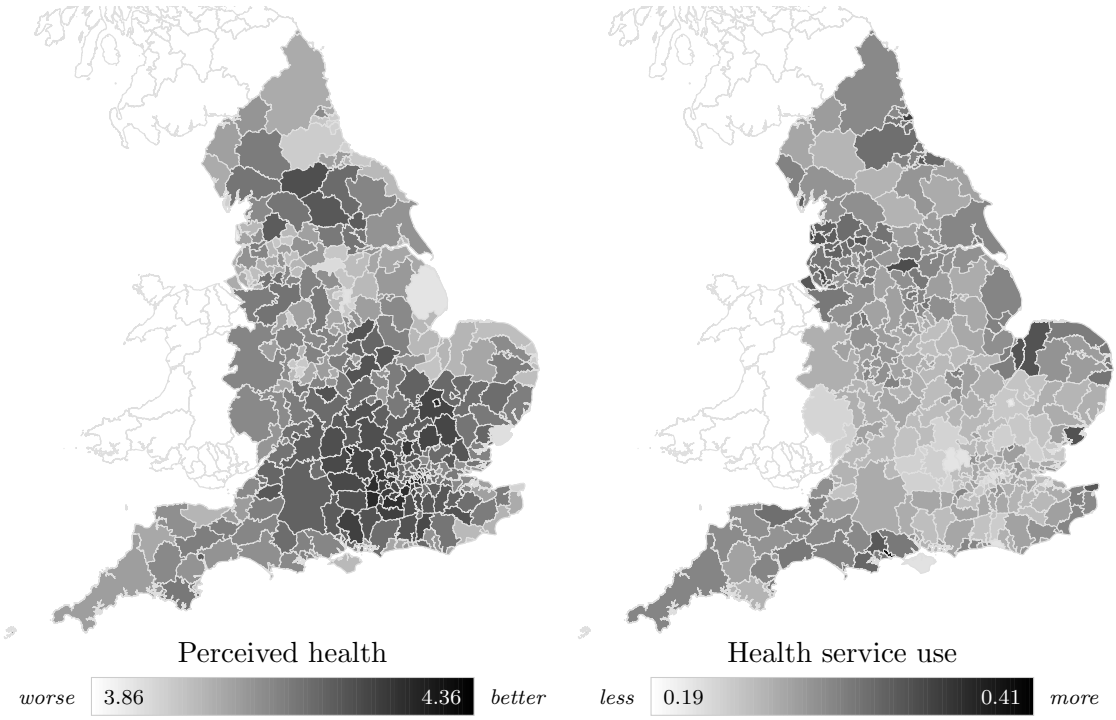


As argued in section 2.4, the social environment is also linked to the geographic area: neighbours, family and friends nearby, and neighbourhood contacts in general remain relevant despite the Internet (Hampton and Wellman, 2003). The role of place for Internet use and health will be explored in the next section – first in terms of area-based and country differences, and then with a short case study of the Isle of Wight as an example for how this research delivers findings down to the local level.

6.3 Geographic environment

Health and Internet use are known to differ by region (NHS England, 2010; Go ON UK, 2015; Blank et al., 2014). Figure 23 shows the aggregate health status and health service use by local authority district. To a certain extent, this map reflects the demographic structure of England – with older people, for example, having lower perceived health and going to hospital more frequently.

Figure 23: Maps of outcome concepts per local authority district



Source: Created based on census and HES data directly; verified with SM data.

The next subsection shows that, based on the simulated quantitative data, some differences exist across areas, regions and urban/rural locations. The qualitative data helps to explain some of these differences.

6.3.1 Area-based differences

Many interviewees attributed area-based influences on their health, as the area was *"less stressed, more relaxed, [...] no noisy neighbours"* (Danielle, East Dorset) and *"affluent [and] well serviced by gyms, parks, good food"* (Hugh, Hammersmith & Fulham), but also because *"the more affluent the area, the less takeaways"* (Brian, Birmingham), and as *"people live longer by the coast"* (Karen, King's Lynn & West Norfolk), whereas it *"takes years off your life to be in a polluted city"* (Harold, Hammersmith & Fulham). Of particular interest in this research are area-based influences on Internet use and health. All Internet-based effects (see Appendix IV, p. 298) were run split up by region, area classification, and urban versus rural. Tables 21–23 show the most interesting results of this exploratory analysis.

First of all, Table 21 shows that the β for health information-finding on perceived health differs across regions: the link is strongest in the North. While the qualitative data corresponds to this finding – the two Internet users from these regions, Miriam and Michael, both suggested that the Internet strongly affected how they perceived their health – it is not entirely clear why this relationship would be strongest in the North. Perhaps it is results from fewer users (Blank et al., 2014) and generally lower levels of health in the North (NHS England, 2010), or a methodological artefact (see subsection 7.5.3). This association is not reflected for health service use, which is mostly reduced in the South (East of England: $\beta = -0.026$; South East: $\beta = -0.022$; South West: $\beta = -0.017$), with $-0.02 < \beta < 0$ for all other regions.

Table 21: Regression on perceived health across regions

	Regions								
	London	South East	South West	East of E.	East Midl.	West Midl.	York-shire	North East	North West
Independent variables	β	β	β	β	β	β	β	β	β
Health info. finding	0.031	−0.015	−0.019	−0.014	0.013	0.018	0.042	0.069	0.063
Age	−0.179	−0.177	−0.217	−0.216	−0.245	−0.193	−0.251	−0.290	−0.240
Gender	0.095	0.081	0.088	0.091	0.124	0.112	0.120	0.130	0.121
Education	0.354	0.283	0.257	0.268	0.348	0.378	0.356	0.387	0.371
NS-SeC	−0.110	−0.06	−0.064	−0.089	−0.069	−0.088	−0.114	−0.110	−0.113
Long-term condition	−0.642	−0.456	−0.455	−0.484	−0.491	−0.409	−0.383	−0.410	−0.396
Adjusted R^2	**0.56**	**0.44**	**0.46**	**0.47**	**0.52**	**0.48**	**0.43**	**0.49**	**0.46**
N *(in million)*	**5.46**	**5.59**	**3.40**	**3.70**	**2.94**	**3.57**	**3.27**	**1.59**	**4.37**
Largest CI									**2**

Beyond health information-finding, there were also differences for the relationship between perceived health and the Internet as a way to reduce loneliness (as reflected in the mechanism "finding companionship", subsection 5.2.6). Table 22 shows that the link between the two is particularly strong in rural areas.

Table 22: Regression on perceived health across urban/rural areas

	Urban/rural area				
	Conurbation	City & town	Town & fringe	Village	Hamlet
Independent variables	β	β	β	β	β
Makes me feel less lonely	0.020	0.035	0.008	−0.014	−0.028
Age	−0.208	−0.206	−0.228	−0.234	−0.221
Gender	0.107	0.095	0.090	0.091	0.097
Education	0.387	0.321	0.321	0.239	0.239
NS-SeC	−0.122	−0.087	−0.115	−0.113	−0.087
Long-term condition	−0.494	−0.449	−0.453	−0.487	−0.502
Adjusted R^2	**0.50**	**0.46**	**0.46**	**0.51**	**0.53**
N *(in million)*	**13.24**	**14.58**	**2.94**	**1.94**	**1.20**
Largest CI					**2**

This finding may be due to the circumstance that "social exclusion is almost entirely an urban problem" (Power, 2000, p. 1). This is reflected in Anthony's argument that the feeling that nobody cared was associated with larger cities:

> *When you're walking around in places like London, you could fall dead on the street, and people just run past you. They just got too much to do.*
>
> – **Anthony, 58, Amber Valley**

In addition, living in certain cities may be more depressing; this may also depend on the economic situation of the wider region (Bernard et al., 2007):

> *You get a lot of people with cancer [...] there was four people down the street with it, and I thought, that's a bit too much. You do get a lot of people around here depressed and things. [...] Middlesbrough as a whole sometimes is a bit of a depressed area, a very depressive town. [...] Sometimes it seems everybody's on a downer because job-wise, there's not a lot of jobs for people to get, and then social security is on your back, you have to have a job, you have to look for somebody, so it's a vicious circle. [...] There should be more help locally for people. I don't think there's enough, education-wise, information for people.*
>
> – **Miriam, 59, Middlesbrough**

At the same time, the relationship between the Internet strengthening local communities and health service use was stronger in urban areas. Table 23 shows that individuals who thought that the Internet made their communities more densely knit made use of health services more frequently.

Table 23: Regression on health service use across urban/rural areas

	Urban/rural area				
	Conurbation	City & town	Town & fringe	Village	Hamlet
Independent variables	β	β	β	β	β
Strengthens local community	0.019	0.007	−0.002	−0.001	0.000
Age	0.431	0.418	0.439	0.461	0.468
Gender	0.229	0.197	0.176	0.138	0.111
Education	−0.044	−0.026	−0.000	−0.012	−0.012
NS-SeC	0.047	0.037	0.043	0.021	0.006
Long-term condition	0.024	0.036	0.030	0.010	−0.000
Adjusted R^2	**0.22**	**0.20**	**0.21**	**0.22**	**0.21**
N *(in million)*	**13.24**	**14.58**	**2.94**	**1.94**	**1.20**
Largest CI					**2**

This is in line with Miriam's observation that closer communities are more likely to get medical help in emergencies (see the "monitoring for emergencies" mechanism, subsection 5.1.4). Supporting Hampton and Wellman's (2003) point about the Internet's importance for local communities, Miriam used her smartphone to communicate with others nearby, and also had face-to-face contact with her direct neighbours:

> *The neighbours are aware – they know I had a minor stroke, so they tend to watch out for me as well. [...] People don't need to know all your business, but there are situations when you are on your own and you need people to look out for you. It is nice that people know things about you.*
>
> **– Miriam, 59, Middlesbrough**

The circumstance that Miriam related increased awareness of neighbours to tangible help may partially explain why the relationship between the Internet decreasing loneliness and perceived health was stronger in cities, as tangible help may also be a source of emotional support, especially when the help is voluntary (Schaefer et al., 1981).

Figure 24 maps the regressions of Internet-based effects on perceived health, confirming the regional and urban/rural differences in Tables 21 and 22.

FIGURE 24: Maps of Internet-based effects regressed on perceived health



NOTE: The maps show the standardised regression coefficients for two Internet-based outcomes on perceived health, controlling for age, gender, education, NS-SeC, long-term conditions. They were created by running 326 separate multiple regressions in parallel, one for each local area district (such as East Dorset or Kensington & Chelsea) in England (see source code in Appendix IX, p. 327).

The results, however, need to be interpreted with care: the MAUP challenge described in subsection 3.4.4 may lead to findings that only result from the chosen level of aggregation (Dark and Bram, 2007; Spielman and Yoo, 2009). In addition, some of the patterns in the maps above – for example between North and South – also stem from the way regions were incorporated as part of the constraints in the SM model (see Appendix V, p. 305). This does not invalidate the results; it simply reflects the fact that OxIS participants in certain areas and regions expressed particular views. However, it should be kept in mind that how the data was created also explains part of the information presented on these maps (see subsection 7.5.3).

In addition, the range of regression coefficients from $\beta = -0.07/-0.10$ to $\beta = 0.14$ shows that the regression coefficients per local authority district form a distribu-

tion around the overall mean. As evidenced throughout the interviews, this does not imply that health information-seeking has negative effects in the South, and positive benefits in the North. The regional differences may rather be a consequence of the different levels of health information-seeking and different levels of self-rated health – they do not necessarily express differences in their relationship.

Finally, some area-based factors mentioned by the interviewees cannot be evaluated through the quantitative data. For example, many individuals in larger cities, particularly in London, explained that the wider choice of hospitals affected their health service use, as did the availability of specialists:

> *My mother's melanoma [was treated] at [a particular hospital] [...] and they're the best in London. I feel that many specialist doctors like to be in London, they're less likely to be in Brighton or [...] in Manchester. But I'm a Londonite, so that's how I feel.*
>
> **– Helen, 43, Hammersmith & Fulham**

On the other hand, those in rural areas rarely complained about the lack of choice, often having several hospitals equidistant from their homes. However, distance to the hospital was brought up as a limiting factor for health service use. For example, Wanda explains how getting to the hospital in winter is difficult for her:

> *The biggest difficulty is getting to the nearest big hospital. I do drive, but I'm not what I would call a confident driver these days, and to get to Swindon – or Bath, which is the other main hospital – is quite a drive. [...] If you got to do it in darkness and bad weather, it's not a lot of fun. So I do worry about that, especially as you get older.*
>
> **– Wanda, 58, Wiltshire**

Analysing distance to hospital – using the coordinates of hospitals (NHS Choices, 2014), and the population weighted centroids of the OA (Census, 2011e) – was carried out in the preparatory phase of this research. However, while distance had an influence on both perceived health and health service use, this difference nearly disappeared once the urban/rural variable was included in the model. This indicates that these differences are less determined by the precise distance, and more by urban versus rural locations, as reflected in Helen's and Wanda's quotes above.

6.3.2 Country-based differences

In addition to differences between areas, individuals also brought up several factors that related to the UK versus other countries. This became evident in subsection 5.1.8, about the Internet affecting satisfaction with health services, with Ian

concluding *"how lucky we are being UK citizens"* (p. 164) in comparison to the US with its different health insurance coverage schemes, for example. In addition, several interviewees brought up differences between US and UK websites (Nettleton et al., 2005), preferring UK websites on the grounds that people are *"much more measured"* (David, p. 167), *"a bit more factual"* and *"not so dramatic"* (Helen).

Interviewees also immediately recognised NHS websites as official and trustworthy based on the parallel with the offline world. This both applied to non-users such as Kathleen, who explained that *"[her] friend in London always uses the NHS pages – she doesn't go to the Internet"*, but also for users such as Hugh:

> *The NHS is actually quite good, it comes quite high on searches if you enter a medical problem. So NHS obviously, it's an amazing brand name for these things, so I'd click on NHS first or something like that. Something which I feel has some kind of non-commercial purpose to it.*
>
> **– Hugh, 45, Hammersmith & Fulham**

Particularly the worriers, who said they preferred health information from official sources, argued that they wanted to see that a website was from England, where they trusted the protection mechanisms more than in other countries:

> *There [...] should be more organised websites, [...] so you feel confident you're going to the right website. I think it needs to be policed more [...] [by] the government. [...] Maybe even websites that don't have the dot-co-dot-uk should not be allowed on, or you should know [...] it's a different company, a different group of people [...] not based in England. Because if it's based in England, then you have more laws to control websites.*
>
> **– Helen, 43, Hammersmith & Fulham**

The interviewees also provided many examples of how "localised" they perceived online discussions and information to be. This may be a result of Google's search algorithms taking location (or at least country-level differences) into account through the user's IP address, and Facebook making contributions from members of one's own social network or a local network more salient even in public groups:

> *I've spoken to someone who's in Singapore, because I said certain creams, we have that in this country – and don't know why, I automatically assumed it was in the UK, and then it wasn't until I looked and went "Oh, actually, it's a different country".*
>
> **– Danielle, 30, East Dorset**

Isabelle also placed more trust in individuals from her country, particularly those from the Isle of Wight, where she grew up and has lived ever since. This highlights again how the Internet serves to connect local groups (Hampton and Wellman, 2003):

> *On Facebook I suppose I feel safer [compared to the Internet in general], because I know that [...] most people are probably from the island, or maybe Britain [...]. It's almost like you trust it more. [...] You just trust other Britons, or you think "I want someone from the Isle of Wight."*
>
> – **Isabelle, 38, Isle of Wight**

The Isle of Wight will be detailed further in the next and last subsection, which introduces a short case study for this specific local authority district.

6.3.3 Case study: Isle of Wight

The case study of the Isle of Wight serves two major purposes. The substantial reason is showing how this research can provide insights down to the local level. A methodological reason is that, as data was integerised for this research partly due to the practical requirement of decreasing the size of the resulting dataset by about a factor of 7 (see subsection 3.2.2), this subsection uses non-integerised data in comparison. The Isle of Wight was chosen because it has urban and rural areas, and because there were four interviewees from this local authority district, who provided several insights related to features of places.

First of all, analysing the access, support and usage divide in Table 24 shows that the general relationships identified in Chapter 4 still hold true. For the access divide, next-generation use is related to higher levels of perceived health ($\beta = 0.062$) and lower levels of health service use ($\beta = -0.038$), although particularly for the latter, the difference from first-generation users ($\beta = -0.024$) is more pronounced than for all of England. The same applies to the support divide, although it diverges from the pattern for all of England at one point: the relationship for unsupported users to both perceived health and health service use is slightly stronger than for supported users ($\beta = 0.041$ and $\beta = 0.033$; $\beta = -0.016$ and $\beta = -0.019$). Finally, the association between health information-seeking and both outcomes remains relatively weak for the Isle of Wight as well ($\beta = 0.012$ and $\beta = -0.005$).

With respect to the difference between integerised and non-integerised data, Table 24 shows that the differences between both are rather small: in most cases $\Delta\beta < 0.01$, with the largest relative deviation below 6%. Though small, this deviation is still important, and confirms in particular that small associations ($\beta < 0.01$) should not be interpreted as going into one or the other direction, as the sign of the coefficient may change through the use of non-integerised data, apart from the circumstance that the error may be larger due to the data being simulated.

Table 24: Regression for access, support and usage divide (Isle of Wight)

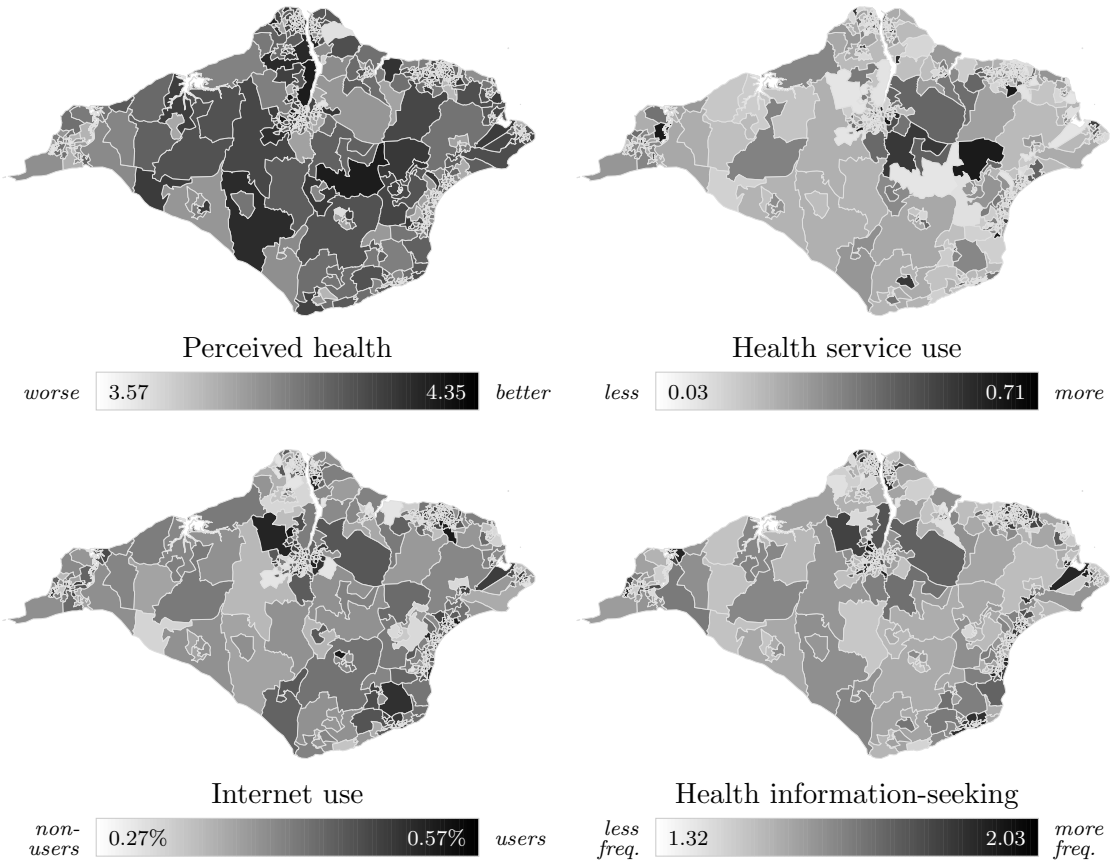
	Perceived health			Health service use		
Independent variables	β	β	β	β	β	β
Next-generation user	0.059			−0.038		
	0.061			−0.039		
First-generation user	0.039			−0.024		
	0.039			−0.022		
Ex-user	0.006			−0.001		
	0.007			−0.002		
Never-user *(omitted)*	–			–		
	–			–		
Independent user		0.062			−0.035	
		0.066			−0.037	
Supported user		0.033			−0.016	
		0.033			−0.015	
Unsupported user		0.041			−0.019	
		0.040			−0.018	
Proxy access		0.017			0.001	
		0.019			−0.001	
Proxy availability		0.009			−0.003	
		0.012			−0.004	
Fully excluded *(omitted)*		–			–	
		–			–	
Health information-seeking			0.012			−0.005
			0.012			−0.002
Age	−0.219	−0.226	−0.003	0.403	0.408	0.005
	−0.222	−0.228	−0.003	−0.405	−0.410	−0.001
Gender	0.105	0.105	0.052	0.116	0.116	0.071
	0.103	0.103	0.051	0.114	−0.114	0.062
Education	0.248	0.248	0.047	−0.010	−0.010	−0.001
	0.262	0.262	0.047	−0.015	−0.016	−0.001
NS-SeC	−0.056	−0.057	−0.014	0.013	0.013	0.005
	−0.061	−0.062	−0.016	0.010	0.011	0.001
Long-term health condition	−0.525	−0.525	−0.363	0.012	0.012	0.015
	−0.523	−0.523	−0.363	0.011	0.011	0.017
Adjusted R^2	**0.60**	**0.60**	**0.50**	**0.20**	**0.20**	**0.21**
	0.61	**0.61**	**0.51**	**0.20**	**0.20**	**0.23**
N *(in thousands)*	**116**	**116**	**85**	**116**	**116**	**85**
	116	**116**	**85**	**116**	**116**	**85**
Largest CI	**4**	**5**	**6**	**4**	**5**	**6**
	3	**5**	**6**	**3**	**5**	**6**

Note: The first line in black provides integerised data. The second line in grey adds non-integerised data for comparison.

In addition, mapping Internet use, health-related Internet use and both health outcomes reveals some patterns (shown in Figure 25): for example, based on the simulated data, health information-seeking may be more prevalent in the more urban areas of the Isle of Wight, particularly around Newport and (East) Cowes, where – according to the interviewees' information – network coverage is better:

> *We used to live [in a very rural area], so you don't get nothing down there. You could get Internet down there, but if you got a mobile phone down there, you got no chance of picking up a signal. So well we've moved [...] obviously and now got Wi-Fi! And it's great, just like look up something or just say "I'll google that" – it's nice to have that.*
>
> **– Isabelle, 38, Isle of Wight**

FIGURE 25: Maps of Internet use and health for OAs on the Isle of Wight



However, the maps – in conjunction with Isabelle's quote above – also highlight some limitations: for example, the simulation does not incorporate available connections and their speed, although broadband may influence some health-related

aspects as well (Whitacre and Brooks, 2014). In this case, the demographic structure may also be different in these areas, so that the map may still deliver valid insights. More generally, however, as Irene recalls from the Skype calls that her daughter on the island set up for her, Internet reception may not be too bad:

> *The picture [on Skype] is quite good. [...] Sometimes the Internet is quite slow. It takes ages to warm up sometimes, it's a bit slow at times. [...] So I wouldn't say it is fast broadband, but people can get it. [...] Reception is not good for some telephones, quite bad in some areas [...] but we're alright, we can use ours here.*
>
> – **Irene, 82, Isle of Wight**

In general, the interviewees said they were happy with the health services available on the island, although average health service use (0.21 = on average, every fifth individual went to hospital once in 2013/14) is lower than the mean of England ($\mu = 0.28$, Table 10). As the map for perceived health shows, health perception is relatively good on average: at 4.00, it is below the mean for England ($\mu = 4.12$, Table 10), although this is not surprising given the sociodemographic structure of the Isle of Wight. For the spatial distribution of perceived health and health service use, it would be worth getting to know the areas in more detail in order to substantiate the spatial findings. According to Ian, areas on the Isle of Wight are very different in terms of the people living there:

> *On the Isle of Wight, [...] if you're working and everything is okay, you get housed with the workers. If you're a bit mentally ill, you get to the mentally ill. If you're a complete retard, you're put up in the retard areas. Obviously if you're a druggy, you go with the druggies. They isolate you, they want to deem this place as a holiday resort.*
>
> – **Ian, 42, Isle of Wight**

Finally, in terms of general living, all interviewees stressed the peace and quietness of the area, which they also associated to positive health effects in comparison to the mainland:

> *It's very relaxing kind of – compared to if you were living over there on the mainland in a city or something like that. It's more hectic there.*
>
> – **Irvin, 85, Isle of Wight**

On the other hand, particularly for more urban areas on the island, isolation was also among the themes that appeared throughout the conversation:

> *I prefer the mainland. I find the water like prison. [...] You can't get on the boat sometimes, and it's extremely expensive.*
>
> – **Irene, 82, Isle of Wight**

> *[The main reason for moving to a more urban area was] isolation. We were very isolated. And also, certain courses that we wanted to go on, it's easier, obviously, being in [an urban area].*
>
> – **Isabelle, 38, Isle of Wight**

This suggests that the association between the Internet as a way to reduce loneliness and perceived health may have positive effects. However, the strength of the relationship on the Isle of Wight ($\beta = 0.003$) closely corresponds to the average for England ($\beta = 0.005$). Then again, this is not surprising, given that the association was stronger in conurbation areas, of which there are none on the Isle of Wight. However, keeping in touch with others far away ($\beta = 0.026$) was more positively related to perceived health than for the rest of England ($\beta = 0.009$). This supports Irene's point about the Internet being an important tool for keeping in touch and reducing loneliness, which she considered a positive influence on health: *"it's not quite like presence, but it's the next best thing."* The resulting implications for policy and practice in relation to the Isle of Wight, as well as methodological implications for the use of SM to derive local insights, will be further discussed in the following chapter (sections 7.4 and 7.5).

This chapter complemented the last missing part of the analytical framework – individual and contextual factors – and showed that both individual and geographic factors largely influence Internet use and health separately. However, the social environment also affected *how* the Internet influenced perceived health and health service use. As a result, all three kinds of influence factors are important to take into account. They will be discussed throughout the next chapter, which discusses the findings from all three results chapters, and thereby integrates them into the wider context of existing empirical literature, theoretical work, in addition to deriving some practical implications.

7 Discussion

As "theory without empirical research is empty", and "empirical research without theory is blind" (Bourdieu, 1988, pp. 774–775), this chapter connects the threads of theory and previous empirical work from Chapter 2 with the findings presented in Chapters 4–6, and the methodology used to generate these insights as described in Chapter 3. This chapter discusses the major cross-cutting themes related to theory, policy and practice, and methodological approaches for understanding the relationship between Internet use and health. Due to the multi-faceted links between these themes across all three RQs and the existence of various emergent findings from the mixed methods design, this chapter discusses the results of this research as an integrated whole, before the conclusions for each RQ and the theoretical, practical and methodological contributions are synthesised in Chapter 8.

The results showed that the analytical framework established in Figures 1–5 is a useful conceptualisation for understanding the relationship between Internet use, perceived health and health service use, but that it needs to be refined and added to. In addition, this research delivers several insights for digital inclusion: first, it proposed an "outcomes divide" in the sense that Internet-based outcomes vary beyond differences in access and usage (DiMaggio and Hargittai, 2001; van Deursen and van Dijk, 2014); second, it showed the interrelated nature of social and digital inclusion, with the social environment being both a catalyst for deriving outcomes, and a potential outcome in itself; third, it argued that some Internet-based mechanisms also exist offline, which indicates that the Internet is evolutionary in the context of health, rather than presenting something entirely new; and fourth,

it found a "lifestyle paradox" expressing the idea that Internet users may perceive their health as lower based on an increased lifestyle awareness, inspired by Albrecht and Devlieger's (1999) disability paradox. Moreover, for the debate on agency and structure as the conceptual frame of this research, the results suggest that the Internet helps to manage different fields in the Bourdieusian (1986) sense, redefines the medical consultation, may shape the wider socioeconomic structure through Internet-based outcomes, and reconfigures the role of spatial structures.

These theoretical insights also have policy-related and practical implications for several areas in which the benefits of digital technologies for the health system are aspired to: enabling outcomes while leaving nobody behind; introducing digital services; and shifting responsibility from the state to the individual. As another practical insight, this research suggests that the definition of medically relevant data should be expanded, the way large-scale data resources are made available for reuse should be reconsidered, and that how data is obtained from individuals in a complex, digitally mediated environment may need to be rethought.

To detail and substantiate these arguments, section 7.1 starts with an analysis of the mechanisms connecting Internet use and health (as is the focus of RQ_2), linking the empirical findings presented in Chapter 5 to the potential mechanisms identified in the literature review (Chapter 2). This section is presented first, as it goes through the main themes that this research confirmed, redefined or brought to the surface, and thereby provides a starting point for all following sections in this chapter. Section 7.2 examines the shifting digital divide from access and usage to outcomes based on the typology of orientations towards health information-seeking (in the context of RQ_1), the relationship with social inclusion (contributing to RQ_3), non-user mechanisms offline, and proposes the "lifestyle paradox" as one of the main emergent findings of this research. Drawing on further insights on individual and contextual factors from Chapter 6, section 7.3 discusses the duality between structure and agency based on Giddens (1984) and Bourdieu (1986) in the context of health and the Internet (delivering insights for RQ_3). In section 7.4, the focus shifts to a practical perspective by discussing implications for policymaking and health service provision along the lines of the aims for health systems outlined above. The final section (7.5) discusses the chosen mixed methods design, reflects on the qualitative and quantitative data collection and analysis, and suggests specific additions to the secondary datasets used in this research.

7.1 Multiple mechanisms between the Internet and health

This research has demonstrated the existence of bidirectional relationships between Internet use and health. On the one hand, health problems and their concomitant circumstances may keep individuals from using the Internet (Dutton and Blank, 2013); health problems may induce health information-seeking and health service use (Dwyer and Liu, 2013; Xiao et al., 2014; Houston and Allison, 2002); and health service use may create or reduce the need for health-related Internet use (Caiata-Zufferey et al., 2010). On the other hand, acknowledging this feedback loop (Andersen, 1995), this research has found multiple ways in which the Internet influences health outcomes – establishing these is the main purpose of this research and they will be discussed in this section.

In terms of causality for the Internet's effects on health, the prerequisites following Hume (Miles and Huberman, 1994, see subsection 3.4.4) – a constant relationship, a plausible mechanism, and cause temporally preceding effect – deserve more exploration. This research has shown that the idea of the relationship constantly holding true may need to be redefined: it may differ between types of health information-seekers (see Chapter 4), in different situations (see Chapter 5) and across individual and contextual circumstances (see Chapter 6). This is in line with the pragmatic paradigm, which – unlike (post-)positivist paradigms – allows for individual differences in the relationships identified, and shows the importance of incorporating antecedent, individual and contextual factors into this research. In terms of plausible explanations, this research found multiple mechanisms influencing perceived health and health service use in different directions, similar to Ziebland and Wyke's (2012) argument that Internet-based outcomes may be both positive and negative. While specific mechanisms do not only exist for certain types of health information-seekers, some manifestations of these mechanisms vary slightly based on the typology, and will be highlighted where applicable.

To match the observed mechanisms described in Chapter 5 with the analytical framework built in Chapter 2, this section systematically maps both against each other. It starts with the Internet's affordances (subsection 7.1.1), discusses to what extent individuals build up resources as part of the process (subsection 7.1.2), and then proceeds to those factors antecedent to health service use and perceived health (subsection 7.1.3). This serves to verify and complement the analytical framework, and lays the ground for further theoretical discussion on digital inclusion (section 7.2) and the relationship between structure and agency (section 7.3).

7.1.1 Affordances

As presented in subsection 2.3.1, the Internet has three main affordances: content (diverse user-generated or official health information online, available anytime and anywhere), connection (contacts formed or fostered online with a wide range of individuals), and mediation (personal unobserved access, reduced or increased cues for communication, and the spanning of geographic distance). These three affordances are highly interrelated and may be perceived jointly by individuals, but are analytically separable for the mechanisms presented in Chapter 5. Figure 26 shows each mechanism mapped to the affordances of the analytical framework. The grid demonstrates that all mechanisms may be attributed to at least one affordance, and that all three affordances are relevant in the context of how Internet use may influence health. This is not self-evident, as the literature on health and the Internet has a predominant focus on health information-seeking (Powell et al., 2005). Since these categories are exhaustive, no further affordance category needs to be added to the analytical framework; the remainder of this subsection compares the findings to the more granular affordances set up in Table 2.

Table 26: Theoretical affordances versus observed mechanisms

	Analytical framework		
Mechanisms	Content	Mediation	Connection
Avoid medical help	■	□	
Encourage service use	■	□	□
Estimate urgency	■	□	
Monitoring for emergencies		■	■
Spurring discussions	■		
Being more demanding	■		
Making decisions about treatments	■		
Affecting service satisfaction	□		
Re-evaluating a situation	■	□	
Comparing/feeling normal	■	□	□
Exchanging best practices	■	□	□
Affecting lifestyle perception	□	□	□
Managing identity/boundaries		■	■
Finding companionship		□	■
Being embedded w. family/friends		□	■
Changing life circumstances	□		

Legend: ■ Major effect □ Tangential effect

Content stands out as a driver of many mechanisms in Figure 26. In line with previous research (Powell et al., 2011; Ziebland and Wyke, 2012), all of the more detailed affordances outlined in Table 2 came to the fore: instant information (Whitney: *"[the main advantage is that] it is instant, because I worry about things in waiting"*, p. 116); information around the clock (Ian: *"[at night –] that's when all the crazy thoughts go through your head"*, p. 175); the active and personalised information-searching (Ian: *"it's a very personal thing, no one can see what I'm doing"*, p. 152); and the diversity of content (Brian: *"You can cross-reference things [...] rather than [depend on] the opinion of one person like the GP"*, p. 137). While the diversity of content was valued – and enabled comparisons of oneself to others, feeling normal, and the ability to tap into user-generated health information – diverse content also caused outcomes not to be achieved. This particularly related to situations with unclear symptoms or potentially severe consequences (subsection 5.1.1), or when individuals found exaggerated or incidental information (subsection 5.2.1). In addition to health problems often being the motivation for health information-seeking (see subsection 4.3.1), this may partially explain why health information-seeking was associated with higher levels of health service use (Dwyer and Liu, 2013; Lee, 2008; Suziedelyte, 2012). However, in line with Alemi et al. (1996) and Gustafson et al. (1999), this research generally found that Internet use and having found something health-related online were related to lower levels of health service use for most individuals (see section 4.5).

Additionally, what this research adds is a more nuanced understanding of information preferences. Whereas other studies concluded that individuals preferred a mixture of personal stories and factual information (Entwistle et al., 2011; France et al., 2011), the typology in this research reveals more about who prefers what: while learners and sceptics saw user-generated information online as a way to learn about other people's stories and crowdsource information, all the other groups highlighted the potential dangers of lay information online and mostly used the Internet for official, factual information. While this may be dependent on the specific condition, as reflected in Whitney's statement that *"basically the people that know most about [MS] are the people who got it"* (p. 119) and in previous research on genetic disorders (Schaffer et al., 2008), the typology in principle cuts across different long-term conditions and thereby presents a valuable systematisation of individuals' information preferences, which will be discussed further in subsection 7.2.1.

Mediation is tangentially related to most mechanisms (see Figure 26). This research also confirmed the majority of affordances: enhanced experience due to

seeing family members as opposed to just hearing their voices (Barbara: *"It would just be on the phone [...] in which you don't have a monitor [...] and it will pick me up [when I am low]"*, p. 178), also due to various forms of information (Miriam, *"it shows you the picture on the Internet [...] all the little instruments used"*, p. 118); cue reduction and asynchronicity (Michael: *"That's how I got through [my mental problems] [...] it is a lot easier to talk to someone online than [...] in person."*, p. 175); mass communication and crowdsourcing (Danielle: *"Then I used Facebook actually and said to [all other] parents 'Any advice?'"*, p. 119); constant communication for emergency monitoring (Nathan: *"Now we have a monitor from the doctor that looks after my health"*, p. 159); and making it possible to span larger geographic distances for keeping in touch with others. In these instances, the Internet provided unique affordances that do not exist offline or with other media.

For **connection**, the Internet allowed both daily coordination with those nearby and ensured help in emergencies (Kevin: *"If something is wrong, I ring them up and [...] they sort it out for me"*, p. 159), for keeping in touch with others far away (Barbara: *"I keep in touch with them on the Internet, [otherwise] I'd probably be on a higher dose of medication [for my depression]"*, p. 178), and finding new ties (Nathan: *"We find it a very good companion [...] it fills a very large void to use this"*, p. 175). This confirms the positive effects on health for communication with family and friends (Bessière et al., 2008), particularly for depression (Bessière et al., 2010; Kraut et al., 2002). In line with Rains and Keating (2011) and Bessière et al. (2008), using the Internet for companionship was mediated by social embeddedness (see subsections 5.2.6 and 5.3.2), while also providing evidence for how individuals may build up affective relationships with the technology itself (Reeves and Nass, 1998; Pols and Moser, 2009). However, for obtaining user-written information from others, prolonged connections were not necessarily valued. Many individuals rather read about others' accounts in forums as a one-off effort, rather than building up a network. This links to the question of whether resources (such as social capital) are built up through these affordances, as will be discussed below.

7.1.2 Resources

Matching the resources from the analytical framework created in Chapter 2 to the observed mechanisms described in Chapter 5, Figure 27 shows that the hypothesised resources play a role, though information and knowledge-based resources are found in more mechanisms. Although the number of mechanisms in the grid does not equate to importance, information and knowledge capital were also particularly

highlighted by the interviewees. In addition, economic and cultural capital were added to the framework. While these are linked to the existing resources in the analytical framework – for example, knowledge capital can also be subsumed in cultural capital (see subsection 2.3.2) – they are added here separately based on the findings of this research, as will be explained in this subsection.

TABLE 27: Theoretical resources versus observed mechanisms

	Analytical framework				**Additions**	
Mechanisms	Information	Knowledge capital	Social support	Social capital	Economic capital	Cultural capital
Avoid medical help	■	■				
Encourage service use	■	▫				
Estimate urgency	■	■				
Monitoring for emergencies			▫	■		
Spurring discussions	■	▫				▫
Being more demanding	■	■				▫
Making decisions about treatments	■	▫			▫	▫
Affecting service satisfaction	■	▫				
Re-evaluating a situation	■	▫	▫	▫		
Comparing/feeling normal	■	■	▫	▫		
Exchanging best practices	■	■		▫		
Affecting lifestyle perception	▫	▫				
Managing identity/boundaries	▫			■		
Finding companionship			▫	▫		
Being embedded w. family/friends			■	▫		
Changing life circumstances		■			■	■

LEGEND: ■ Major effect ▫ Tangential effect

Information played a role in many contexts: above all, for evaluating the need for professional help (Wanda: *"It would help me to decide whether I actually need to go and see a doctor, or whether [...] [it] can be treated otherwise"*, p. 117); creating legitimacy for help-seeking (Ian: *"The Internet [...] made me feel as though I am normal, I am entitled as a human to [...] question this situation that is embarrassing"*, p. 151); changing the medical consultation (Martha: *"I [...] look it up and [could] be a bit more demanding if I knew it could be a sign of something"*, p. 162); understanding the wider context of the health system (Ian: *"I was watching this documentary online [and realised] how lucky we are being UK citizens"*, p. 164); but also for re-evaluating a situation and feeling normal about it (Miriam: *"It just*

made me feel like 'stop whining about it and live your life.' These people are a lot worse off", p. 133); and learning health-related tips from others.

In most cases, it is more appropriate to speak of information only – in the sense of a one-off fact in the specific situation. However, particularly for the learners, and even more so for the learners with a long-term health condition, the Internet provided something that can be referred to as **knowledge capital** (Savage, 2013), and enabled them to become experts in the condition (Ziebland, 2004). This particularly applied to those mechanisms related to prolonged interaction, such as avoiding medical help and estimating urgency (Miriam: *"I read it on the Internet [...] [and] so I was more aware that it wasn't this urgent"*, p. 158), being more demanding (Danielle: *"now I know that other people are trying natural remedies, I want to try these remedies"*, p. 163), feeling normal, but also changing life circumstances more generally (Michael: *"half the things I learned on healthcare is actually from the Internet"*, p. 180). Interestingly, this sense of "becoming an expert" was not the case for everyone with a long-term condition: for example, Matthew, a pragmatist, Kevin and Susan, adigitals, or Helen, a worrier, did not display the same sense of curiosity about wanting to know and building up long-term knowledge capital.

Social support was mainly obtained either from known sources, for example through being monitored for emergencies (Miriam, p. 159) or by being embedded with family and friends, where in particular the Internet's mediation affordances (such as constant communication, or spanning geographic distances) came into play. In addition, individuals experienced health-relevant support from others online, as reflected in Ian's example of *"an extra friend without the need of getting rained on or spending patrol in your car"* (p. 176) that he could connect to without mentioning his mental health issues. Interestingly, individuals also described social support-related phenomena through reading others' stories online (without necessarily connecting to them), particularly through comparing themselves to others and, as a result, feeling normal about their situation. This may link to earlier findings that the mere perception of support helped individuals feel better (Schaefer et al., 1981; Wethington and Kessler, 1986) – with no actual support being provided.

For **social capital**, some exchange of user-generated information was enabled only by long-lasting connections, for example Danielle's *"mini mums-net"* (p. 119) or Ahmad's network for exchanging weight loss tips (p. 170). In Danielle's case, the Internet primarily served to amplify the reach of communication, thereby supplementing face-to-face contact and neither increasing nor decreasing social capital (Wellman et al., 2001). In Ahmad's case, his network not only helped him to

exchange tips with other likeminded individuals, but also generated praise and emotional support (p. 170). Then again, most individuals interested in user-written information, including those with long-term conditions, trawled through different forums to read others' stories (as a one-off effort) rather than building up networks, which would represent a form of social capital. It rather seemed that individuals used the Internet as an ad hoc network, a phenomenon also cited among the motivations for health-related Internet use (Powell et al., 2011; Ybarra and Suman, 2006). At the same time, this underlines that the Internet, and technology more generally, often affect healthcare practices in unforeseen ways (Pols and Willems, 2011): while individuals primarily looked for information, they also obtained social support in the sense of feeling normal about their situation or feeling part of a bigger group. These aspects, as well as the potentially negative aspects of Internet-based social support that Sarah (p. 179) and Susan (p. 176) brought up, will be discussed further in subsection 7.2.2 in the context of social and digital inclusion.

Beyond the capitals derived from the earlier literature, the findings suggested that Bourdieu's (1986) other capitals are relevant as well: economic capital as an asset that is "immediately and directly convertible into money" and cultural capital as something that is "convertible, on certain conditions, into economic capital and may be institutionalized in the forms of educational qualifications" (Bourdieu, 1986, p. 47). **Economic capital** through content obtained online became evident, for example, in Matthew's statement on finding out about benefits after an accident, which would *"make all the difference to whether you make it or not"* (p. 179). Ian's example of the Internet as a tool for learning *"long lavish words that one can use in job interviews"* shows how **cultural capital** can be converted into economic capital. Beyond job interviews, learning words to know what to say in consultations and appear more knowledgeable (Ian, p. 152) may be seen as cultural capital. Similarly, Mary's attempt to find out about her ailment so she could react more knowledgeably to the doctor's diagnosis (p. 160) is another instance of building up small portions of cultural capital, and expressing it through a changed habitus in the consultation.

These links to perceived health and health service use can also be supported theoretically based on Bourdieu's capital conversion in conjunction with Grossman's (1972) health capital. As discussed in the literature review, viewing health as a capital good that depreciates over time suggests that individuals may invest in their health in order to have time for other activities. Grossman (1972, p. 241) suggests that "the higher a person's wage rate, the greater the value to him of an increase in healthy time," so that higher economic capital may be associated with increased

health. A similar logic also relates to health service use: based on the premise that individuals primarily want good health, health service use is essentially a means to obtain it (Folland et al., 2014). As individuals experience higher education and wages (or economic and cultural capital, following Bourdieu, 1986), the opportunity costs of using health services increase as well – as reflected in Hugh's statement of the loss of working time (p. 134).

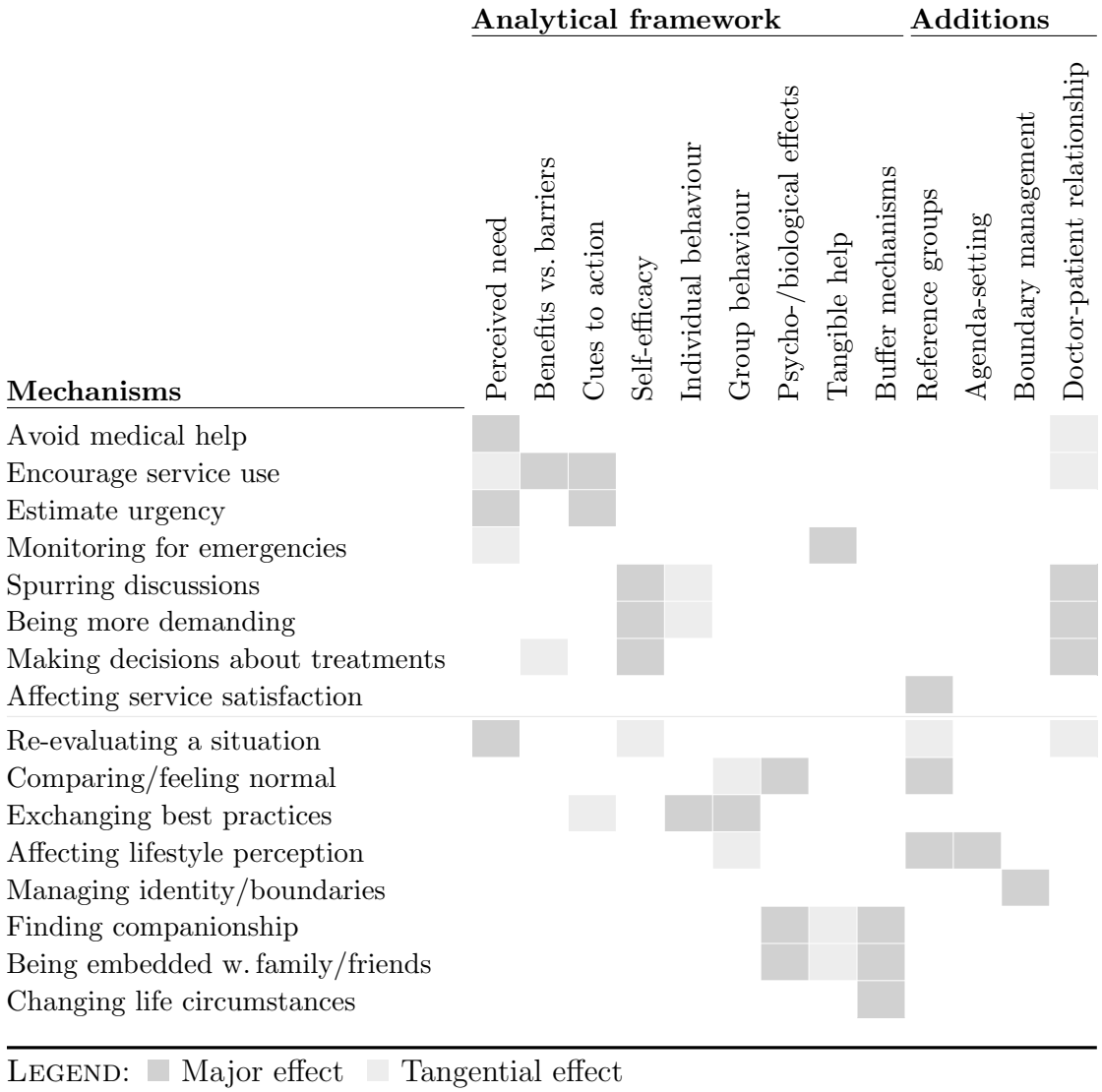
This relationship between economic/cultural capital and health is also reflected in Tables 17 and 18, showing that higher education and socioeconomic status are associated with higher levels of perceived health and lower levels of health service use. Similarly, the quantitatively modelled mechanism around the Internet affecting life circumstances (Figure 17) shows mostly positive links between the Internet's effects and perceived health. Consequently, although Bourdieu (1986) referred to the embodied state of cultural capital as "*long-lasting* dispositions of the mind and body" (Bourdieu, 1986, p. 84, emphasis added), altering short-term habitus in the consultation may help individuals to feel that they are being taken more seriously. This aspect will be discussed again in section 7.3.

7.1.3 Health factors

Finally, Figure 28 shows that the theoretical factors influencing perceived health and health service use identified in Chapter 2 exist in at least one of the observed mechanisms described in Chapter 5. Interestingly, while the analytical framework distinguished between effects on health service use and perceived health, these differences are less clear-cut in Figure 28, supporting the interrelated nature of both health outcomes, despite their relatively low covariance throughout the analyses. Then again, some observed mechanisms cannot be matched to any of the previously identified factors, and some relate to more aspects than those covered in the analytical framework. Consequently, the analytical framework is complemented by four factors: reference groups; agenda-setting; boundary management; and the doctor-patient relationship. This subsection discusses these new factors, and those previously identified as part of the analytical framework.

Perceived need was influenced by many observed mechanisms, particularly those that related to avoiding medical help, estimating urgency and re-evaluating situations (Wanda: *"[I] saw photographs similar, and sort of hand and heart made the decision not to follow it up"*, p. 155). Changes to perceived need were also visible for encouraging health service use, but this mechanism above all related to changing the balance between expected **benefits and barriers** for seeking

Table 28: Theoretical health factors versus observed mechanisms



Legend: Major effect Tangential effect

help, both mentally and through learning concrete ways to address certain topics (Ian: *"now I can make a GP appointment, [and] know what to say because I had references from the Internet"*, p. 152), which jointly with the quantitative results in section 5.3 reflects Berger et al.'s (2005) findings about health information-seeking for stigmatised conditions. In many cases, the Internet also provided specific **cues to action**, primarily through explicit messages found online about help-seeking (Danielle: *"If it says 'nothing to worry about' or 'leave it 48 hours,' then I just follow the advice that is given"*, p. 158) and how quickly to seek help (Miriam: *"on the Internet it says 'you have to be at the hospital within 4 hours'"*, p. 158),

sometimes put into action with others' help (Mary: *[My family] would look it up [...] and let me know what it sounds like, and then I'd say I'd go to the doctor's*, p. 208). This shows how both information and the social environment shape health service use (Goldsmith et al., 1988; Pescosolido, 1992; Dingwall, 2001, see subsection 2.1.1).

While these factors decreased health service use for many interviewees, the Internet had no effect for others. This was mainly the case when health problems were new and unknown, with potentially severe consequences, and particularly when symptoms could not be described clearly (see subsection 5.1.1). In these cases, individuals nevertheless went to see a professional. The worriers and sceptics also provided instances in which the Internet actually increased their health service use, particularly due to an increased perceived need and cues to action obtained online. Finally, just as Ian (p. 157) and Haythornthwaite and Wellman (2002) highlighted, Internet experience played a crucial role in how people interpreted the cues generated by factual or user-written information online, as also reflected in the quantitative data for health service use (see subsection 5.3.3).

At the same time, the Internet increased **self-efficacy** (Bandura, 1986, 2004), thereby enabling individuals to ease their complaints by taking action themselves, with implications for both health service use and perceived health. In some cases, these related to **individual behaviour** changes, such as Brian's example of drinking water (p. 136) or Miriam's example of buying different shoes due to her arthritis (p. 171). However, most examples of self-efficacy related to a changed doctor-patient relationship, both in the sense of spurring discussions (Helen: *"I was on that computer the whole time looking up the specialists so we could ask questions, because you only have 20 minutes"*, p. 117), being more demanding (Martha, p. 162), and making decisions about treatments even as a non-user (Kathleen, p. 141).

While the **psycho-/biological effects** are not detailed in this research, several interviewees attributed positive effects on perceived health to the Internet for connections with friends and family, and new contacts to cope with struggles in life (Michael, p. 175). This also related to tangible help in the sense of being cared for (Schaefer et al., 1981), particularly in the context of monitoring for emergencies (Kevin, p. 159). For getting help in emergency situations, the boundaries between the Internet and traditional mobiles were particularly blurry, as the mobile already fulfilled many benefits that individuals described. These benefits particularly came to the fore for those with long-term conditions: for example, Barbara thought the Internet was particularly helpful for her depression as she was able to see her family members on the screen (p. 178), similarly to the particular relevance of tangible help

elements for those who had experienced strokes (Kevin and Miriam, p. 159). These effects on coping with the circumstances created by health conditions therefore also suggest a certain **buffer effect** (Cohen, 1988).

Group behaviour influenced individual behaviours in the sense of getting new ideas from others online, as in Danielle's examples about seeing a doctor for natural remedies for her child's eczema (p. 163). Moreover, individuals incorporated ideas about food and fitness in their daily routines, in addition to positive praise (Ahmad: *"It makes me feel good when they say 'Oh, you lost so much weight'"*, p. 170). However, while the previous literature often regarded nutrition and exercise as positive group influences, and promoting smoking or alcohol as negative influences on health, this research showed that perceived health may also be decreased as a result of guilt felt about not leading a healthy lifestyle. Even though this is taking the definition of a group relatively far, individuals may become part of a wider and less clearly defined group of people online (Wellman and Hampton, 1999).

The idea of an enlarged group for comparisons leads to the first addition to the analytical framework: **reference groups**. Merton and Lazarsfeld (1950) described how individuals do not judge themselves on an absolute scale, but rather relative to others. This insight has been described in the health literature before (Buunk et al., 2013), and has been confirmed in the context of the Internet, for example when women viewed pictures of other women online (Tiggemann and McGill, 2004). While individuals tend to compare themselves to those that they aspire to (Merton and Lazarsfeld, 1950), this research showed that there is both upward and downward comparison (similar to Festinger's (1954) social comparison theory about opinion formation). Individuals compared themselves "upward", giving them hope that they could *"come out on the other end"* (Miriam, p. 133), but also "downward" by finding examples of individuals who had more severe health problems or higher weight (Helen, p. 169). Given the breadth of experiences expressed online, finding someone worse off (with a particular health condition) may be easier than it is offline. Reference groups will be revisited in subsection 7.2.4 on the "lifestyle paradox", and again in subsection 7.3.4 on the Internet-related reconfiguration of spatial structures.

Related to changed reference groups online, and leading to the second addition to the analytical framework, one may say that the Internet fulfils a certain **agenda-setting** function. Agenda-setting describes the idea that the media do not necessarily decide what individuals think, but that they determine what they think *about* (McCombs and Shaw, 1972). Consequently, the Internet may influence lifestyle perception, making it difficult to *"avoid the kind of media awareness around*

lifestyle" (Hugh, p. 173). Agenda-setting may be an unintended effect of the Internet at an individual and societal level (Brown and Walsh-Childers, 2009), and may come from an interplay of social media, advertisements, the promotion of health and wellness apps, and wearable technology, in the sense that these suggest certain ideal lifestyles that individuals should aspire to. As this research has shown, this has moved the bar against which individuals rate their own health. The "lifestyle paradox" will be detailed further in subsection 7.2.4.

Third, this research suggested that individuals' health perception is also influenced by their **boundary management** of different parts of their life – either through upholding a healthy identity after a stroke vis-à-vis professional contacts despite being speech-impaired (Kevin, p. 174) or through *"join[ing] as a normal character without the mental health"* (Ian, p. 152) online. While the literature has primarily discussed the Internet as a connecting technology, enabling a "shift from living in 'little boxes' to living in networked societies" (Wellman and Hampton, 1999, p. 648), it may also help to create boundaries between different parts of life. This observation illustrates the well-established finding that "on the Internet many people recast identity in terms of multiple windows and parallel lives" (Turkle, 1997, p. 72). However, it is less understood that this still applies today, with the online and offline division increasingly becoming deconstructed (Orgad, 2009); the Internet's mediation affordance also serves as a boundary even between offline spheres, and this may positively affect individuals' health perception.

Finally, while the **doctor-patient relationship** as a frequently researched outcome (Powell et al., 2003; Broom, 2005; Anderson et al., 2003; McMullan, 2006) was not the focus of this research, it was indivisible from perceived health and health service use in the interviews, and should therefore be integrated into the analytical framework as a mediator. Spurring discussions, being more demanding, and making decisions about treatments had implications for seeking professional help (Ian: *"now I can make a GP appointment [...] because I had references from the Internet"*, p. 152), for ameliorating concerns about procedures (Miriam: *"because I've read it up exactly what they would do [...] I was fine"*, p. 162), and for coping with ailments that could not be confirmed in medical settings (Michael: *"when I go to the doctors, [...] then there's nothing there [so] I just go for it online"*, p. 169), apart from a heightened sense of self-efficacy with positive effects on perceived health.

These additions to the analytical framework (reference groups, agenda-setting, boundary management, doctor-patient relationship) are revisited as part of the debates on digital inclusion (section 7.2) and agency versus structure (section 7.3).

7.2 Delineating the shifting digital divide

This research addresses several aspects of the digital divide. First, it supports the claim that the digital divide is shifting (DiMaggio and Hargittai, 2001; van Deursen and van Dijk, 2014; van Deursen and Helsper, 2015) – while the access and usage divides still remain relevant, this research also suggests an "outcomes divide" in subsection 7.2.1. Another important factor is the social environment (Warschauer, 2004), which – as this research proposes – acts both as a catalyst for outcomes, and is an outcome in itself. At the same time, inequalities due to different levels of support from the social environment persist (see subsection 7.2.2). Subsection 7.2.3 relates back to the mechanisms and discusses to what extent the way in which the Internet affects perceived health and health service use also applies offline, asking to what extent the Internet presents something novel or simply reinforces existing trends (see literature review, subsection 2.2.2). Finally, as touched upon in the last section, subsection 7.2.4 discusses the "lifestyle paradox" similar to Albrecht and Devlieger's (1999) disability paradox, in the sense that the Internet creates some lifestyle awareness that may negatively affect health perception.

7.2.1 Towards an outcomes divide

In line with the shifting digital divide from the access to the usage divide (DiMaggio and Hargittai, 2001; van Deursen and van Dijk, 2014, see subsection 2.2.1), Helsper (2012) and van Deursen and Helsper (2015) have argued for unequal outcomes, as resources built up online only partially translate into individuals' "offline" lives due to several factors around attitudes, skills, and content found online. This research supports this shift to an "outcomes divide" or "outcomes inequalities"[11], as it showed that aspects around Internet access, support and usage shape its outcomes (support in particular will be discussed in subsection 7.2.2). However, this research also found that understanding Internet-based health outcomes requires a more nuanced view.

Based on the results of this research, the typology of health information-seekers is useful to conceptualise these outcome inequalities, which transcend the boundaries of the traditional sociodemographic dimensions of the digital divide (van Deursen and van Dijk, 2014, see subsection 4.4.1). As described in sections 4.3 and 4.4, the typology is set up based on antecedent factors around motivations and enablers, which confirms Hale's (2013) finding that previously existing health behaviours

[11] In line with DiMaggio and Hargittai's (2001) argument that there is not a single digital divide, one may rather speak of inequalities

translate to the online realm, as well as Helsper's (2012) emphasis on attitudes and skills as mediators for Internet outcomes. For example, being a user in the learner or pragmatist group was associated with higher perceived health than for other types after controlling for sociodemographic factors. In support of this, the qualitative data suggested that learners and pragmatists used the Internet efficiently in health-related contexts, both to increase the appropriateness of their health service use and, for the learners, to gain self-efficacy for self-care and to extend professional advice.

This does not imply that outcomes are positive. For example, the medical sceptics had a strong relationship between health information-seeking and finding something that improved their health, and reduced their health service use accordingly, whereas the relationship to perceived health was negative (see Table 15 and Figure 14). This may indicate that low-trust individuals tend to substitute physician services with health information online (Dwyer and Liu, 2013; Lee and Hornik, 2009), but not necessarily with positive health effects. For the worriers, finding something health-related online was barely associated with any changes in perceived health and health service use. This suggests that they do not necessarily use health service use more frequently, as occasionally evidenced (Helen, p. 139), but also that this group barely obtained any Internet-based health benefits.

Finally, the effects of Internet use were lowest for delegators and adigitals, although with higher effect sizes for the former (see Table 15 and Figure 14). As this research has shown, some of the delegators used the Internet more intensely than some of the users, building support networks with various individuals for different purposes. While both the delegators and the adigitals largely comprise non-users, this shows how Internet-based outcomes may not follow the lines of the user/non-user split. This is further corroborated by the relatively similar quantitative results for users and non-users in the delegator group, which may partially be due to their preference for outsourcing health information-seeking.

With the typology, this research represents the first systematisation of orientations towards health information and associated outcomes, in addition to classifications of general orientations and patient roles in digital health systems (Lemire et al., 2008; Tjora, 2014). Like other typologies of general Internet use (Dutton and Blank, 2013; Christodoulides et al., 2013) which also cover the range from enthusiastic users (such as the learners) to less eager users (such as the adigitals), the developed clusters "are overlapping, but each has a distinctive profile" (Dutton and Blank, 2013, p. 4). As described in Chapter 4, individuals are assigned the most suitable type reflecting their orientation towards health information-seeking

and its effects on the two health outcomes analysed in this research, although they may not necessarily only exhibit characteristics of one type. For example, Isabelle, a learner, reported an incident where she *"worried more, for no reason"* as she read about a virus infection that might endanger her father's health (p. 166), while Anthony, a worrier, displayed adigital tendencies as a result of being concerned (*"It's too much information, or too much wrong information. So I don't tend to look on it now"*, p. 149). Moreover, based on the lack of health variables in OxIS for building the quantitative foundation of the typology, 6 out of 31 OxIS participants had to be manually reclassified (see subsection 4.4.1), suggesting that the typology's dimensions do not fully capture the *health-related* motivations and enablers yet.

In addition, the typology reflects orientations towards online health information, and as with political orientations, these may change (Jennings and Niemi, 1975). Consequently, individuals may change types over the course of their lives: for example, Ian may have exhibited different characteristics before his move, and highlighted the role of experience (p. 157), while Elizabeth attributed her medical scepticism to her experiences with her own stroke and her sister's gastric cancer (p. 163). This is also a corollary of the disease process models introduced in subsection 2.1.1: just as dynamic models may be needed which recognise health and help-seeking "as part of a broader, socially embedded, 'illness career', comprising the complex processes, multiple 'small' decisions and 'recipes' of action" (Wyke et al., 2013, p. 80) and take into account individual trajectories of health (Gately et al., 2007), so may these events change some of the antecedent factors around motivations and enablers upon which the typology is based. This is also shown in Andersen's (1995) classification of antecedent factors having medium to high mutability, and the central role of experience for online activities and outcomes (Haythornthwaite and Wellman, 2002). While this research is only cross-sectional and thereby cannot assess movements between types over time, the fact that some individuals needed to be reclassified based on their changed or specific attitudes about health-related Internet use (see subsection 4.4.1) supports the somewhat permeable nature of the typology, even though the general proportions of attitudes towards the Internet have remained remarkably stable over the last decade (Dutton and Blank, 2013).

Ultimately, the typology presents a tool for systematising health information-seeking, a way to conceptualise the outcomes divide, and a basis for formulating practical implications. This is similar to other typologies in the health psychology domain (Clatworthy et al., 2005), including the above-mentioned cultures of the Internet, which serve to "shift study of the Internet away from an overly narrow

focus on comparing users and non-users, and [focus] more research and debate on other variations among users and non-users that have equally significant implications for the future of the Internet" (Dutton and Blank, 2013, p. 9). Although clustering and segmentation have their origins in market research (Punj and Stewart, 1983; Christodoulides et al., 2013), typologies can deliver both theoretical value – for example, proposing the outcomes divide in this research – and practical value (Doty and Glick, 1994; Clatworthy et al., 2005). In particular, they can close the gap between singular theoretical findings and practical implications (Zubin, 1938), as will be detailed in subsection 7.4.1 on the practical implications of this research.

7.2.2 Social and digital inclusion

Related to Internet-based outcomes is the relationship between social and digital inclusion, as "the inequality that does exist is social, not digital" (Warschauer, 2003, p. 297) and the socially excluded are also the digitally excluded (Helsper, 2008, see subsection 2.2.1). In the context of health, this research delivers two main insights for the link between social and digital exclusion. First, social inclusion may help in use of the technology, and act as a catalyst for health-related outcomes. Second, the Internet may also reduce social exclusion through its affordances, making benefits around social support an outcome in itself, with potential benefits for health (Cohen, 1988; Berkman et al., 2000; Uchino, 2004). Both will be discussed in this subsection.

The catalyst role of the social environment is crucial as support may be necessary to find and interpret health information, for example due to a lack of skills (Shirley, p. 120; Karen, p. 121), the device not being accessible enough (Karen, p. 108) or being too small (Miriam; Wanda, p. 121), or because the individuals lacks their own device (Barbara). This particularly related to older females, who may have lower technical skills based on their age, while women generally tend to underestimate their skills (van Deursen et al., 2011; Hargittai and Shafer, 2006, see subsection 7.5.2). Several of the interviewees mentioned above explained how their social environment helped them to use the Internet and achieve outcomes – with the exception of Shirley, who, however, also showed low interest in health-related Internet use (see section 6.2). In addition, Ethan described how his medical friends helped him with *"successive approximation"* (p. 207), evidencing how individuals overcome a lack of information-finding skills and health literacy through their social environment (Eynon and Malmberg, 2012; Hargittai, 2002; Lee et al., 2004; Bakardjieva, 2005). This form of support has also been termed apomediation, which – as opposed to intermediation – is about lending a hand to help others "navigate through the

onslaught of information afforded by networked digital media" (Eysenbach, 2008, no page). Particularly for the worriers, the social environment also allows them to reflect on information found online, as *"you have to be really in a supportive environment to be able to look up on the Internet some things"* (Helen, p. 138).

However, intermediation still plays an important role for non-users' Internet access, for example via their carers (Sadasivam et al., 2013), and as a social information filter (Chapple et al., 2012). In intermediation, the credibility and quality of the information depends on who looks something up (Eysenbach, 2008), which can also bring substantial advantages to the individual. For example, as Helen explained, *"if there was anything really negative and really harmful, then [she] was able to siphon that out"* (p. 208), similar to Isabelle's example of not telling her father about what she had read online (p. 166). The delegators in particular appreciated this filtering, either to avoid negative concerns (Mary: *"When I read things, I take it as the worst thing ever [...] I'm like 'oh no, that's bad, that's bad,' and then I'm sitting there panicking"*, p. 147) or to benefit from other people's skills in terms of technical access and information-seeking (Norman, p. 111; Kathleen, p. 141). This again highlights that some people do not wish to use the Internet due to the convenience of proxy use (Reisdorf et al., 2012). Even though non-users may not obtain the benefits related to the Internet's mediation affordance and actual use, this research showed that health information obtained through other channels may still be beneficial.

At the same time, social inclusion may also be an outcome of Internet use itself, with effects particularly on perceived health, either directly or through social support derived from the network (Cohen, 1988; Berkman et al., 2000; Uchino, 2004). For example, related to the mediating affordance of the Internet, 90-year-old Nathan referred to the technology as a *"very useful companion"* and *"filling a large void"* (p. 175), thereby seeing the technology itself as providing company – showing "that there is no opposition between cold technology and warm human care" (Pols and Moser, 2009, p. 175) – and as a medium for connecting with his family. As 83-year-old Irene explained, loneliness may be one of the ways by which health deteriorates at later stages in life (p. 178), as empirically found in biomedical research as well (for example, Holt-Lunstad et al., 2010; Cacioppo et al., 2002).

In line with early research on the potentially isolating effects of Internet use (Kraut et al., 1998, 2002), some individuals also reported negative effects of the Internet, as *"that's the problem with society today: most people spend most time hidden up indoors, instead of getting out there and getting a life"* (Sarah, p. 179), and *"people don't talk to people anymore"* (Susan, p. 179). In contrast to Kraut et al.

(2002) and Bessière et al. (2008), these negative effects were not necessarily explained by the existing social network (see subsection 5.3.2), but rather were a feature of the adigital type – although the direction of causality may be disputed. Particularly in cities (where social networks are more loosely defined) the relationship between Internet use leading to less loneliness and perceived health was strongly positive (see subsection 6.3.1). This supports the importance that Kraut et al. (2002) and Bessière et al. (2008) attribute to the existing, offline social network.

Finally, some mechanisms also rely on previously existing social networks: for example, Danielle's reference to her Facebook contacts as her *"mini mums.net"* (p. 119) shows how online and offline social networks converge, and illustrates how the Internet's mediation affordance allowed for the crowdsourcing of advice from people with whom trusted relationships have existed before. As another instance of converging online and offline realms, Miyata (2002) concluded that receiving support online also promoted giving and receiving help offline. This stands in contrast to boundary setting online (see subsections 7.1.3 and 7.3.1), although the two do not invalidate each other: while the crowdsourcing example mentioned above reflects a higher degree of openness for less personal topics (where individuals even used their real names, see p. 119), boundary management applied particularly for embarrassing or debilitating conditions, as with mental health or strokes.

7.2.3 Non-users' mechanisms

If some of the Internet's affordances also exist offline (see subsection 7.1.1), and outcomes are shaped by factors antecedent to health service use (see subsection 7.2.1), then some Internet-based outcomes may not be exclusive to the online realm. Several non-users presented ways in which health information – similar to the Internet's content affordance – played a role in their health service use and perceived health. Natalie, for example, would use her book on diabetes, and discuss advice in there with her doctor, because *"obviously sometimes the book can be misleading [...] [and] you just want to check before you follow that information"* (p. 161). Her desire to ensure accuracy follows a similar discourse as in the context of the Internet (Eysenbach et al., 2002). Similarly, Kathleen (who had a positive image of doctors and the Internet) also explained that her doctor *"doesn't just do it like from what it says in the paper [...] but takes it home and googles it on his computer"* (p. 141), expressing how her doctor ensured the quality of newspaper content.

Similarly, for the connection affordance, Charlotte argued that talking about potentially embarrassing conditions such as incontinence alleviates the issue – *"half*

of the problem goes away, just by talking about it and turning it into a joke and swapping ideas with everybody how to cope with it" (p. 169) – although Charlotte did not think of the Internet as a suitable medium for doing so. In addition, the ways in which non-users perceived their health was evaluated relative to a neighbour (Catherine, p. 106) rather than to others. This shows the different reference groups for non-users, as will be touched upon again for the lifestyle paradox (subsection 7.2.4) and the reconfiguration of spatial structures (subsection 7.3.4).

These "non-user mechanisms" lead to three points about the Internet and the media landscape. The first relates to whether the Internet creates *new* influences on health. As this subsection showed, this is not the case for those mechanisms enabled by the content and connection affordances, as these exist offline in similar ways. In these contexts, the Internet may have created evolutionary rather than revolutionary change (Powell et al., 2011) and led to new ways of doing old things (Gershuny, 1983). The Internet certainly facilitates information-finding, as locating information offline is considerably more time-consuming and difficult in terms of access for most patients (Broom, 2005; Anderson et al., 2003), and also enables the identifying of others with similar conditions to talk to (Wright and Bell, 2003). The existence of these mechanisms for non-users also reflects the central role of antecedent factors around motivations and enablers (see sections 2.2 and 4.3).

The second implication is that other media are not necessarily seen as more trustworthy. Whereas the prevalent picture seems to be that non-users do not trust the Internet, and rather rely on newspapers, television and the radio (Dutton and Blank, 2013), this research shows a more nuanced picture. In particular, it may correspond to MacKenzie's (1999) "certainty trough", which is characteristic of online trust (Dutton and Shepherd, 2006). In the context of health information, the certainty trough says that both those alienated from the Internet, as well as the "experts", consider health information online to be less trustworthy. This may explain why the adigitals (such as Charlotte and Shirley), who have limited exposure to the Internet, particularly dismiss online health information, while some frequent users, such as the worriers, also do not necessarily embrace health information online – whereas delegators with some indirect exposure (such as Kathleen) even regard the Internet as a way to verify content from newspapers (p. 141).

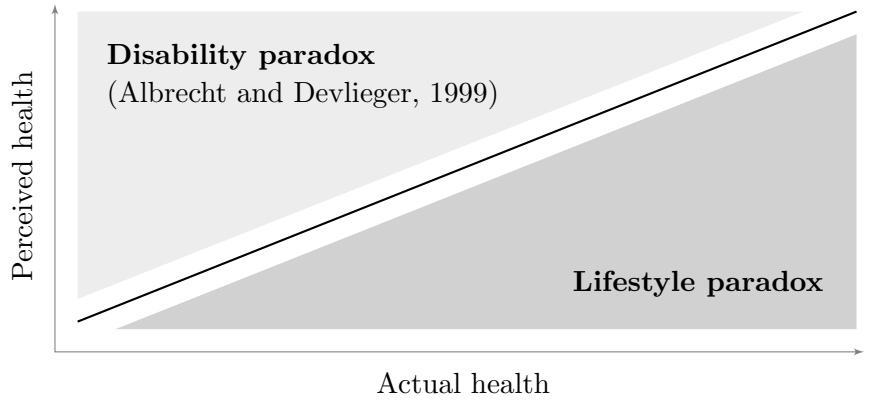
Finally, the non-user mechanisms show the central role of physicians as "sounding boards" for many individuals – either to check advice provided online or in a book, or to use the doctor as an intermediary for accessing the technology. Just as many researchers concluded that Internet use may complement health service use rather

than replace it (Powell et al., 2011; Lupiáñez et al., 2013), this indicates that many individuals use information to work with their doctor (Ayers and Kronenfeld, 2007), with the caveat that the qualitative interviewees did not include any non-user of the medical sceptics type, for whom this may potentially look different.

7.2.4 A new lifestyle paradox?

Finally, the results of this research suggest a certain "lifestyle paradox". This suggests that how individuals perceive their health may decrease based on using the Internet, as the Internet creates a lifestyle awareness that shifts the yardsticks against which individuals rate their health. The idea is inspired by Albrecht and Devlieger's (1999) disability paradox, which captures the finding that individuals with severe health problems may still rate their health as very good, as their baseline for perfect health is lowered. Figure 29 shows a conceptual overview of both paradoxes as deviations from the generally positive association between perceived and actual health (Goldstein et al., 1984; Hunt et al., 1980; Fillenbaum, 1979).

Figure 29: Conceptual overview of lifestyle paradox



The lifestyle paradox is evidenced by three main sources in this research. First, it was evident in how individuals rated their health, often referring to 4 rather than 5 (on a scale from 1–5) because of only being *"kind of averagely fit"* (Nicolas), considering it overconfident to say 5 (Hugh), or *"sometimes [not doing] as much [as possible] fitness-wise"* (Martha, p. 106). Second, it was visible in the quantitative data: Internet use, Internet importance and health information-seeking showed the strongest correlation to perceived health for the rating of 4, although, particularly for health information-seeking, the causality between Internet use and perceived health may rather run the other way round. While this research did not obtain discrete

ratings from 1 to 5, the analysis was made possible by partitioning the perceived health outcome values (see subsection 3.2.2 and Appendix VI, p. 306) into quintiles and by running a multinomial logistic regression on each value (see section 5.3). However, a limitation on this is that the ratings do not exactly correspond to the box that individuals ticked. Third, the interviewees also explained the underlying rationale for this relationship: the Internet may create *"pressure on fitness, health clubs, diet, the whole kind of lifestyle angle that the Internet is very filled with"* so that *"you can't avoid the kind of media awareness around lifestyle,"* which leads to *"a guilt if you're not having that lifestyle"* (Hugh, p. 173).

Supporting the empirical findings in this research, previous research by Whitacre and Brooks (2014) found that broadband adoption displayed a significant positive association with perceived health rated as "good health" (3), but was significantly negatively related to "very good health" (4), and not associated with either "excellent health" (5) or "poor health" (1). The authors do not have a clear explanation for their finding, and hypothesised that it was "perhaps due to a realisation of all of the different aspects that health can entail", and "that more information available online allows individuals to have a more realistic assessment of their own health situation, at the intermediate levels" (Whitacre and Brooks, 2014, p. 774). The results are not entirely comparable due to the slightly shifted scale (with 2 out of 5 as the neutral option), but still support the finding that the relationship between Internet use and perceived health may not be linear.

Challenging the idea of an Internet-based lifestyle paradox, Blaxter (1990, p. 33) states that a common observation across younger people is that "many who described health in terms of fitness appeared to be saying guiltily [...] 'I am not as fit as I should be, therefore – even if I have no illness – I am not healthy.'" The qualitative data confirmed this observation, as primarily younger individuals assumed the fitness view of perceived health (Brian, aged 54, was the oldest to do so). In addition, the external validation of the SM data showed that females in the age group of 18–24 rated their health as lower than the next age group (25–34), with no differences between users and non-users (p. 314). Then again, in support of the Internet-based lifestyle paradox, the quantitative analysis in subsection 5.3.1 showed that there is a strong monotonic link between age and perceived health, and the qualitative data clearly evidenced a link between Internet use and lower levels of perceived health. Finally, Hugh argued that *"the Internet has magnified [and] amplified that dialogue"* (p. 173), so that the Internet may have primarily reinforced existing tendencies among younger people to rate their health as lower based on not being fit enough.

As touched upon in subsection 7.1.3, Merton's (1957) reference groups may provide a theoretical explanation for this phenomenon, as individuals compare themselves to those they aspire to. In this case, this relates to those more engaged in healthy behaviours – but as individuals may choose to represent only certain parts of their identity online, or stress certain aspects of their lives (Turkle, 1997; Mallan, 2009), individuals compare themselves to a group whose lives they know based only on selected aspects. In addition, the lifestyle paradox may be explained by some level of agenda-setting online (McCombs and Shaw, 1972), as advertisements online and social networks make individuals think about lifestyle, and convey that they can and need to shape their health similar to other aspects in life (Lewis, 2006). In contrast to traditional agenda-setting effects, in which the public agenda is set by the news media, this highlights the development of agenda-setting towards distributed networks (McCombs, 2005). However, the topics in the mainstream media and online do not differ greatly, so that the Internet may still have a certain agenda-setting function despite diversified authors and channels (Wallsten, 2007; Lee, 2007).

In that sense, the development of the lifestyle paradox is embedded in a general tendency towards the individual assuming more responsibility for his or her own health (Lewis, 2006), which also entails a growing importance for fitness as a way of preventing chronic conditions and premature mortality (Stevens et al., 2002). The shift towards more individual responsibility is both envisioned by governments (Department of Health, 2010, 2012c) and may be seen as characteristic of post-modern societies in general (Giddens, 1991). Of course, communication media may have intensified this trend and are fundamentally interwoven with larger societal transformations (Thompson, 1995), which provides further support for the reinforcing role of the Internet in the context of the Internet-based lifestyle paradox.

The lifestyle paradox also links to the duality of agency and structure, as the online realm may provide structures that both enable (by generating certain outcomes, see subsection 7.2.1) and constrain (by creating by-products, such as the lifestyle paradox) individual agency. The next section discusses four other aspects related to the interplay between individual and collective forces, all of which contribute to understanding the association of Internet use and health.

7.3 The duality of agency and structure

The analytical framework included factors related to agency (such as motivations) and structure (such as enablers, and individual and contextual factors). As will be

discussed in this section, both of these interrelated forces shape each other, and are relevant in the health context. They affect the approach to Internet use and health information-seeking, attitudes about health service use and health professionals, and ultimately the main focus of this research: Internet-based health outcomes.

Agency has a central function for Internet-based health outcomes, as nobody "has yet invented the drug or written the law that will make individuals do what those individuals must choose to do themselves" (Mirowsky and Ross, 2003, p. 28). This research showed that if individuals are not *interested* in general and health-related Internet use, they will not obtain outcomes – as opposed to those with a high level of agency, but who experience structural constraints (such as devices) that have prevented them from going online. Agency may also shape structure, for example through improving general life circumstances. Then again, structure can both restrain and enable agency (Giddens, 1984): social structures are important enablers for Internet-based outcomes, whereas agency is also constrained by socioeconomic and educational structures. These structures shaped the individuals' sense of being responsible towards the health system, their appreciation of the medical profession, and their ability to find and interpret information (see section 6.1).

These structural constraints relate back to Bourdieu's (1986) concept of fields. While Bourdieu does not provide examples of fields, several non-mutually exclusive field boundaries may be defined in the context of this research – for example, the field of the medical consultation. For health, the link between agency and structure has frequently been researched under the umbrella of patient empowerment (Greener, 2008). However, it is more than that: individuals consciously or subconsciously used the Internet to manage fields in their lives, thereby shaping sociodemographic structures, whereas the Internet also reconfigured spatial structures. This section shows that the interplay between agency and structure is a useful conceptualisation for understanding the links between Internet and health, and includes Bourdieu's concepts of field and habitus as valuable concepts in this context.

7.3.1 Creating and managing fields

The first argument is that with the Internet, individuals are able to create different fields without the structural constraints they would otherwise face. This is primarily enabled by the Internet's mediation affordance, and particularly influenced health perception, but also their health, as shown in the three examples below.

First, the Internet helped create new fields online, where social positions are not yet determined and different structures apply. For example, Ian described the

Internet as a place to be *"a normal character without the mental health"* (p. 152), and therefore to occupy a position without stigma. This supports Bessière et al.'s (2010) finding that social use in particular improved depression: while Ian still rated his mental health as 2 only, he would have considered it as 0 in the past (p. 106). This point is not without controversy, as much of the literature argues that online and offline lives are becoming enmeshed, with the Internet "descend[ing] from the firmament and becom[ing] embedded in everyday life" (Haythornthwaite and Wellman, 2002, p. 4) – and indeed, individuals asking medical questions on their Facebook status update (Danielle, p. 119) or joining support groups with their real name (Isabelle, p. 170) supports this trend. Ian's case shows that the Internet may still be a separate field with different structural properties, with self-conception as an important influence on health perception (Thoits, 2013).

Then again, the mediation affordance also plays a role when real names are used. For example, the Internet may help individuals to maintain a position within a field despite a decrease in resources: although Kevin could no longer speak properly after his stroke, he used channels with reduced cues, such as texting, to keep up his professional network. He may not necessarily return to work, so that the possibility of deriving economic capital from the network may be minimal. However, keeping up his position in the field gave him a sense of a possible future. This optimistic outlook and feeling of control over this outlook are a possible theoretical link through which his perceived and actual health were affected, as empirically shown and theoretically argued in research on self-efficacy and the sense of coherence (Schneider et al., 2004; Suominen et al., 2001; Bandura, 2004). Although Kevin already hated computers before his stroke, this shows how mobile devices with messaging apps may fulfil the needs of someone with very debilitating long-term conditions.

Finally, the last example particularly relates to managing fields. Following Bourdieu (1986), individuals reproduce structures of a particular field into their habitus, for example as habits and conventions. Barbara feels that whenever she is around her private social circle, she is implicitly or explicitly forced into drinking (*"because if I see them at their place, we usually have a little social drink"*, p. 174), whereas she was simultaneously trying to fight her alcohol addiction. While this confirms the negative link between group behaviours and individual behaviours (Reblin and Uchino, 2008; Lindström et al., 2000), it shows how the Internet can also affect (actual and perceived) health by inhibiting certain mechanisms that exist in everyday life, such as the social pressure to drink alcohol.

7.3.2 Levelling the field in the consultation

At the same time, the Internet may also form the field of the medical consultation. There exists a fairly established literature about how the Internet may shape the doctor-patient relationship (Anderson et al., 2003; Broom, 2005), in conjunction with the doctor-patient role sets (Merton, 1957) generally developing towards more active patient roles over the last decades (Gray, 2002). In addition, the medical consultation has often been connected to patient empowerment, in the sense that the power distribution has shifted, and patients have obtained more control over the consultation (Roberts, 1999). Although the consultation itself was not necessarily an outcome of interest in this research, asking individuals about the Internet's effects on health and health service use showed that the doctor-patient relationship is an inseparable element of this, and so will therefore also be discussed in this subsection.

The findings supported the view that the Internet helped some individuals to become more informed in advance of the consultation, ask more questions, and feel better prepared before meeting a medical professional (Broom, 2005). This applied to all types of health information-seekers except for the adigitals, although the depth of inquiry differed (with learners and worriers preparing more detailed questions). However, in line with previous research, many individuals (particularly the adigitals) did not aspire to play this active part in their health and make decisions about treatments (Powell and Boden, 2012), so that the Internet had no effects for them.

In addition, only few individuals felt they could openly talk to their doctor about health information online, and used their knowledge tacitly or when the need arose – although only a few had actually experienced negative reactions. The tendency to use knowledge based on the specific demands of the situation is similar to what Bourdieu (1990) described as the "logic of practice" or as "in-the-game", "where strategies are not predetermined, but emerge and operate according to specific demands of action and movement in time" (as cited in Barrett, 2014, p. 4). Two strategies in the context of agency and structure, and habitus as the way through which structures are internalised in individuals' practices, are discussed below.

The first is the use of keywords to make a smarter impression in the consultation – or to pretend to have higher levels of cultural capital through a consciously modified habitus. For example, the Internet helped Ian to learn *"long lavish words"* (p. 179) so that he *"went in, made [...] an appointment, used keywords from [...] the Internet, [but] never mentioned the Internet"* (p. 152). Knowing these words enabled Ian to bring up the topic in the first place, but he also used them strategically to get himself taken more seriously by the doctor. The explicit use of signalling words reflects

Bourdieu's and Giddens' notion that although habitus conveys individuals' positions in a field, they can change their appearance. Although Bourdieu (1977) sees habitus as "durably installed generative principles" (p. 78) that shape individuals' practices, he also concedes that habitus does not determine action, and that it is possible for individuals to alter their habitus if they are aware of it (Bourdieu, 1984). The strategic use of keywords in the consultation is even more in line with Giddens (1984), who argues that individuals are reflexive agents who may consciously alter their appearance to change their position in a social structure.

A second relevant instance is that of Mary, who asked her family members to look up health information for her (although she had Internet access herself), and then *"relate[s] some of it back to the doctor, like 'my family's read this and says it's this'"* (p. 208). While she did not explicitly say that her primary aim was to be able to refer to others, it is interesting that she uses her family instead of saying that she looked it up online (and may have misinterpreted the advice). In that sense, she assumes some agency in her health, but does not assume responsibility for her findings. Then again, she stated that her health information was for her to be able to say *"Oh, 'it's not that bad,' or 'I know what I got to do,' 'I've got to do this'"* (p. 160). This again shows parallels to Ian's example, showing how Mary consciously alters her reactions to appear more knowledgeable to the doctor.

In addition, the development towards questioning medical authority (Gray, 2002) in combination with health information online may have also led to more individuals challenging medical expert knowledge (Hardey, 1999), as reflected in the medical sceptics type. However, there may be limits to medical lay knowledge, particularly when it leads to the questioning of conventional treatments (Prior, 2003; Broom, 2005). The weaker relationship between Internet use and perceived health, as well as the lower levels of health service use for the medical sceptics, reflects findings from a time before the Internet was widely available: those sceptical towards medical care had lower health service use, but also worse health behaviours and a lower perceived health status (Fiscella et al., 1998; Sutherland and Verhoef, 1994). As the medical sceptics were of slightly lower socioeconomic status and education (Table 13, Fiscella et al., 1998, Sutherland and Verhoef, 1994), one may also argue that sociodemographic structures shaped their attitudes towards health services.

Finally, apart from the medical sceptics, individuals valued the structures available to them in the doctor-patient relationship (Ayers and Kronenfeld, 2007). As Giddens (1984) highlights, structures are not necessarily constraining, but also fulfil an enabling function: some saw the doctor as a sounding board, while others wanted

to follow the expert advice as closely as possible with the help of information obtained online (see subsection 7.2.3). The different mindsets of the learners, pragmatists, worriers, delegators and adigitals versus the sceptics therefore reflect two of Lemire et al.'s (2008) different logics of patient empowerment through the Internet: the professional logic in the sense of executing the prescribed way as well as possible, and the consumer logic about making decisions based on one's own judgement.

To summarise, the Internet has changed the medical consultation: it has redefined the structure of this field in the sense that the Internet contributes to changing individuals' habitus, or to their ability to pretend to have higher levels of cultural capital.

7.3.3 Shaping sociodemographic structure

As discussed in subsection 7.1.2, the Internet may also shape wider sociodemographic structures through building up economic and cultural capital: Internet use entailed benefits such as doing better at job interviews, getting benefits after accidents, and learning online in order to pursue a medical career (see subsection 5.2.8). This confirms Eynon et al.'s (2016, forthcoming) longitudinal research, which found that Internet use may positively affect social mobility over time, thereby shaping individuals' long-term structures with benefits for health, as individuals gain feelings of control and self-worth as their circumstances change (Bottero, 2004).

Two interrelated questions remain in the context of agency and structure. The first is the magnitude of the Internet's impact, with generally low mutability of sociodemographic characteristics (Andersen, 1995). The quantitative data indicated an aggregate effect on perceived health, and the qualitative data provided various instances in which the Internet made a difference in terms of building up economic and cultural capital – but is this impact meaningful? The interviewees indicated that it was: for Matthew, the Internet made *"a difference to whether you manage"* (p. 179) in the context of his work-related accident (although the Internet here did not change his socioeconomic status, but rather helped to sustain it despite hardship). Similarly, if Michael *"didn't read up on the topics, then [he] would not be doing [the new career path] at all"* (p. 180). In that sense, the Internet did make significant and meaningful changes to individuals' socioeconomic status.

The second question is to what extent these changes affect health. In principle, it is well established that higher education and socioeconomic status are related to better health (Adler and Newman, 2002; Chandola and Jenkins, 2014). The first explanation from economics is that health has a higher economic value for

those with higher education and socioeconomic status (Grossman, 1972). This may directly influence health, as it does not assume lifestyle changes, which may be transformed more slowly. The second explanation is around different lifestyles and health behaviours typical of certain sociodemographic groups (Adler and Newman, 2002). If these are changed due to the Internet, it is likely that Internet use, mediated by these sociodemographic changes, also influences health.

7.3.4 Reconfiguring spatial structures

Finally, spatial structures may be reconfigured by the Internet, with implications for perceived health and health service use. In this context, Bourdieu's (1986) and Giddens' (1984) definitions of fields or locales of interaction are helpful again: they conceptualise individuals in structures, in which they have resources and social contacts shaping their practice dependent on the field/place.

A first observation is that Internet users and non-users employed different reference groups for judging their health. Internet users, particularly next-generation users, often referred to other people they had read about online, whereas first-generation users and non-users tended to compare themselves to others in their area: Karen explained she was much fitter than people around, Catherine noticed that *"the lady next door is in a worse state than [she was]"* (p. 106), and Wanda concluded that many nearby friends *"are older than me, and a lot fitter – so that's what I'm comparing myself to"*. This confirms Bernard et al.'s (2007) point that neighbourhood resources play a more important role for the place-bound population, although Bernard et al. primarily referred to children and the elderly. With the general finding that the area influences Internet and health separately, rather than shaping their relationship, this shows how the Internet reconfigures spatial patterns of who people compare themselves to when making relative judgments about their health. However, as captured by the lifestyle paradox (see subsection 7.2.4), this general finding does not necessarily mean that non-users are worse off – just that location is more important for their health perception, even though non-users in fact have a lower health status (see section 4.2).

A second observation is that as the Internet may allow *"zooming to places that kind of promote certain values"* (Hugh, p. 173), spatial structures may be reconfigured, as briefly touched upon in subsection 7.2.4. Similarly, Wellman and Hampton (1999) argued that the Internet created a shift away from "little boxes" that kept interaction apart both socially and physically. The media more generally may create a shift to replace physical with informational settings, so that individuals' spatial

context is flattened and they increasingly have "no sense of place" (Meyrowitz, 1985, part of the title) – even though the author particularly was referring to the television at that time. Finally, Giddens (1991) argued for the decoupling of time and space in modernity, so that locales are shaped by distant influences, and individuals' actions increasingly become decoupled from clearly bounded places.

A third observation related to spatial structures is that despite the Internet's global nature, many individuals still perceive the Internet as rather local, particularly when it comes to trusting the sources. Many individuals highlighted their preference for British sites on the grounds that these are more down-to-earth, even preferring someone from the local area (Isabelle: *"You just trust other Britons, or you think 'I want someone from the Isle of Wight'"*, p. 218). Although particularly in terms of feeling normal, the global reach of the Internet was highlighted (Ian: *"the world's seen it before"*, p. 145), this shows how the local neighbourhood remains important despite global connectivity (Hampton and Wellman, 2003). This argument will be extended as part of the practical recommendations in section 7.4, where the interplay between individual and collective agency may be important for improving health through behavioural change (Bandura, 2004).

In that sense, the Internet may change the importance of spatial structures for users in how they assess their health, and change how places more generally are perceived. However, spatial structures remain important for individuals' local connectivity, and shapes individual and social contextual factors that each individually affect perceived health and health service use (see section 2.4 and Chapter 6).

7.4 Implications for practice and policy

Beyond theoretical advances, this research intends to have a "measurable social and economic impact on policy and practice", as aspired to in the ESRC's (2015a, p. 16) and MRC's (2014) strategic aims. This includes implications for policymakers, but also for health practitioners. While some findings may apply more generally, this research relates to England only, as the analysis was based on specific circumstances such as free access to healthcare, and several interviewees explicitly referred to "Britishness" in justifying their positions, explaining their preferences for sources of health information, their use of NHS services, and their attitudes towards the Internet (see subsection 6.3.2).

Contributing to the aim of the NHS to increase the population's health while reducing healthcare spending (see section 1.2), the recommendations provided are

centred around three themes: enabling outcomes while improving digital inclusion to ensure that "we do not leave anyone behind in [the] move to a digital by default approach" (UK Government, 2013, no page) and introducing new digital services "to transform outcomes for patients and citizens" (NIB, 2014, p. 1) in subsection 7.4.1; integrating individuals into decision-making, and shifting responsibility away from the state to the individual and their local communities (Department of Health, 2010, 2012c), including knowing about local health needs (NHS England, 2014a) to reduce "unwarranted variation to increase value and improve quality" (NHS England, 2010, title page) in subsection 7.4.2; and enabling data-driven insights for the areas mentioned above, which have implications for how medical data is made available for research (Nuffield Council on Bioethics, 2015) in subsection 7.4.3. Appendix XX (p. 361) contains the two-page policy summary that was sent to decision-makers at NHS England, Public Health England and the European Commission.

7.4.1 Enabling outcomes while leaving nobody behind

As this research has showed, Internet use affects perceived health and health service use, even though these may differ between individuals (see Chapter 4) and may be obtained through different mechanisms (see Chapter 5), which showcases the importance of the Internet in healthcare (Powell et al., 2011; Chaudhry et al., 2006).

However, these outcomes are not uniform across individuals, and some may be left behind (UK Government, 2013). Across contexts, several contradictory policy recommendations have already been designed to narrow digital divides (Yu, 2006). The UK digital strategy (UK Government, 2013) distinguishes between users and non-users, and differentiates users based on whether they have used government transactions or information (not necessarily health-related), while non-users are offline and may either be willing or unwilling to get online. However, as the UK Government (2013, no page) acknowledges, "we need to understand who can use digital services, who can't, and what else we need to provide for people who aren't online." As this research has evidenced, it is not only about "can" (enablers), but also about "want" (motivations) – reflecting previous findings that individuals do not necessarily want to be engaged in their health (Powell and Boden, 2012; Henwood et al., 2003; Department of Health, 2012a).

The typology helps to establish possible approaches (see summary in Table 25). Particularly for the learners, pragmatists, worriers and the sceptics, non-users should be provided with access to digital resources, as most individuals of these types cited the non-availability of devices as the main reason for non-use (see Table 13). In

addition, while neither of these types should be seen as superior or inferior, worriers may benefit from additional guidance: they highly appreciated medical professionals, and suggested that doctors should recommend specific websites (Helen, p. 139), reflecting their preference for professional guidance and managing uncertainty (Fergus, 2013). This stands in contrast to the medical sceptics, who – independently of the Internet – might benefit from building up trust in doctors and the medical system to change their health behaviours, evaluate the appropriateness of health service use, and ensure compliance with medical recommendations (Fiscella et al., 1998; Lee and Hornik, 2009).

Table 25: Practical implications based on typology

	Potential points to address		
Type	*Digital inclusion*	*Digital services*	*Medical consultation*
Learner	Ensure availability of access devices	Provide trustworthiness certifications for online resources (incl. apps)	
Pragmatist			
Sceptic			Address attitude towards medical system
Worrier		Provide trustworthiness certifications for online resources (incl. apps)	Provide specific advice on online resources to use
Delegator	Ensure social inclusion for Internet access		
Adigital	Address motivations for (health-related) use	Address image of online services (not inferior to offline)	
General	Continue programmes on skills that enable contact with online resources and services		Provide "sounding board" Ensure Internet skills of physician/medical staff

Particularly for delegators, but also for all others with lower skills, the social environment is of crucial importance for Internet-based outcomes (see subsection 7.2.2). As the name implies, obtaining value from the Internet depends on being able to delegate – and, as Kathleen (p. 141) and Norman (p. 111) vividly described,

delegators had very elaborate networks for doing so. Nevertheless, based on the quantitative data, the group of delegators also contains 22% of individuals who do not know anyone who could access the Internet for them (Table 13). Here, addressing social and digital inclusion become a joint policy priority, where improving social inclusion would also ensure that "we will not leave anyone behind" (UK Government, 2013) – a wording which nicely captures that not everyone necessarily has to be online, but rather that everyone should benefit from online resources.

Finally, for the adigitals, it may be most important to address motivations for Internet use. This applies to users and non-users alike: 80% of non-users of this type have referred to a lack of interest as the main reason for non-use (Table 13), whereas users often showed critical attitudes towards the Internet and explained why the Internet was not suitable at all for health-related matters.

These type-based findings have practical implications, particularly for England's digital strategy and the NHS efforts to enable digital health benefits (UK Government, 2013; NIB, 2014). Existing initiatives such as the "Widening Digital Participation" project are useful to support enabling factors around devices and skills for finding and evaluating information. However, they are only part of the solution, as they neglect the need for individuals to be interested and to want to participate actively in their health. The typology may be used to understand individuals' orientations towards health information, and to define appropriate pathways for each type based on the recommendations laid out in this subsection. Some actions may not be targeted towards particular individuals, but may primarily be of value to particular types: for example, clearer labelling of government official photos – or developing trustworthiness certifications for apps, as is currently being done by the NHS – may be particularly helpful for the worriers, who preferred "official" information and pictures (Helen, p. 140), and for the pragmatists, who tended to be careful with user-written information.

Similarly, practitioners should be aware of these different types, as this may help in the consultation, as Zubin (1938, p. 509) argued several decades ago: "On the one hand, we have a group of abstract formulations about the relationship between variables emanating from research laboratories. On the other hand, we have individual case histories written by the practitioner. [...] It may be that the middle ground will partake of the nature of a grouping [...] that is, the research workers will come to grips with the single individual in the form of a group of like-minded or like-structured individuals." This is not to say that practitioners should not be interested in individual patients anymore, but rather that the typology may be a

device for understanding individuals' expectations towards practitioners. Across all types, individuals wished that they could discuss information found online with the doctor without being *"scoffed at [...] in a sort of derogatory way"* (Wanda, p. 161). Even more so, individuals desired advice from the practitioners on online resources (Helen, p. 139), which entails the corollary that physicians themselves must be at the forefront of digital technologies (Weber and Entry, 2004).

Finally, while digital health services were not the focus of this research, the findings also deliver insights supporting the aim to partially provide consultations with medical professionals online, for example (NIB, 2014). Particularly the pragmatists, the medical sceptics and the learners valued the instrumental efficiency of online services, and individuals across all types (except for the adigitals) acknowledged the contribution that technology may make in the healthcare sector (for example, Hugh, p. 164; Irene, p. 178). While many people already use digital services that involve trust, such as shopping or banking (Dutton and Blank, 2013) and increasingly expect "to access services quickly and conveniently, at times and in ways that suit them" (UK Government, 2013, no page), the adigitals particularly perceived digital services as inferior compared to face-to-face ones (Sarah: *"[The doctor] gets paid to do a job, he gets paid to come and see you [offline]"*, p. 200). This may again be a question of changing attitudes towards digital services, potentially with the help of medical professionals. Of course, this again relies on medical professionals championing digital services themselves (Weber and Entry, 2004).

7.4.2 Shifting responsibility from the state to the individual

Extending the argument that health-related benefits from Internet use depend on being able to and wanting to engage online, this subsection discusses the broader development of shifting responsibility from the state to the individual and its implications for the relationship between Internet use and health. The NHS has declared this an explicit aim (Department of Health, 2010) – both in the sense of helping those with long-term conditions to become experts in their conditions (Taylor and Bury, 2007), and of encouraging the population to assume preventative behaviours (Greener, 2008). This shift is theoretically reflected in a societal shift towards more responsibility for one's own health in modernity (Giddens, 1991; Cockerham, 2005) and the central role of self-efficacy in health and health promotion (Bandura, 2004).

Through the relevance of antecedent factors around motivations and enablers, this research showed that individual responsibility is key to understanding the health-related outcomes of Internet use. Not everybody wants to take over re-

sponsibility for their health decisions (Henwood et al., 2003; Department of Health, 2010), and there are several limits to the view of the active, interested and potentially empowered information-seeker (Powell and Boden, 2012), partially because the wide availability of health information online may create a "felt imperative to be (or to present oneself as) an expert and critical patient" (Ziebland, 2004, p. 1783). This research has confirmed that the desire to be involved – either online or offline – differs between individuals, based in particular on their socioeconomic and educational backgrounds (see subsections 6.1.2 and 6.1.3). In this research, for example, Mary explicitly asked her family to look things up, delegating responsibility, whereas Sarah put forward the argument that it was not her responsibility to look up health information online (p. 200). Similarly, individuals' sense of responsibility towards making the NHS more efficient, and the potential role of technology in it (see Hugh, p. 164) differed along socioeconomic and educational lines.

However, structures nevertheless remain crucial. The health system continues to play an central role, and the Internet may complement rather than replace existing services for most patients. In the context of the Expert Patient Programme, as an example of another approach that puts forward a high degree of responsibility, Greener (2008, p. 287) argued that "patients, with the experience and knowledge they have acquired about their conditions, clearly can make a positive contribution toward their own care, but this must be tempered with a recognition of the dependency that we place upon doctors – dependency in a positive, caring sense." In addition, better incentives are required for those without intrinsic motivations to be involved in their health – for example, in the form of financial incentives, which may work particularly well with groups of lower socioeconomic status. This has been observed for the interview compensation in this research (see section 3.3), although it must be acknowledged that financial encouragement to take part in interviews may not translate to this type of long-term health behaviour change, as some individuals may not be able to or prefer not to be involved in their healthcare (Henwood et al., 2003; Powell and Boden, 2012).

Finally, social and geographic structures may play an important role in the assumption of responsibility and in shaping health behaviours (Bandura, 2004), in line with Miriam's conclusion that *"there should be more help locally for people"* (p. 213). Section 6.3 showed that links from health information-finding and social outcomes (such as the Internet reducing loneliness and strengthening local communities) to both health outcomes differed between areas. The main qualitative explanations for these geographic differences focused on the choice of health service providers

and distance to health facilities, but also on social cohesion in these areas. As the social environment may also shape the relationship between Internet use and health, and "the approaches that work best promote community self-help" (Bandura, 2004, p. 162), the local level should be strengthened as well. This relates to helping individuals to access digital resources and socially integrating them, as identified as important mediators for Internet-based effects in this and previous research (Bessière et al., 2008; Rains and Keating, 2011).

In addition, the case study of the Isle of Wight serves to derive further local insights in the context of Internet use and health. For example, health information-seeking was found to be more prevalent around Newport and (East) Cowes (see Figure 25). This quantitative finding may be explained by the interviewees' accounts that there was better mobile network coverage and faster Internet access in these areas. As first-generation Internet use was more strongly related to perceived health and health service use for the Isle of Wight than for the rest of England, this implies that the Isle of Wight might benefit from further advancement of their mobile networks and broadband infrastructure. This is also supported by the quantitative finding that keeping in touch with others was more strongly related to perceived health compared to the rest of England, and the qualitative accounts of the importance of the Internet for social contacts. Video-based communication in particular may be important when living on an island, with greater hurdles for regular face-to-face contacts with friends and family members on the mainland.

7.4.3 Allowing data-driven insights for health and healthcare

The insights described throughout this chapter are based on the qualitative and quantitative data presented in Chapters 4–6. This subsection presents some observations related to obtaining the secondary data used in this research, in addition to specific recommendations for all secondary datasets discussed in subsection 7.5.4.

This research has added to the growing body of literature that extends the boundaries of what may be considered "medical research". Whether something counts as health data "depends on the *use to which they are put* as much as the source from which they were obtained or the purpose for which they were originally collected" (Nuffield Council on Bioethics, 2015, p. 19, emphasis added). In this spirit, there has been an appreciation throughout the last decades that medically relevant research extends beyond the confines of healthcare facilities, and has to incorporate sociodemographic characteristics (Marmot and Wilkinson, 2006; London Health Observatory, 2010; British Academy, 2014), lifestyle and health

behaviours (Cockerham, 2005; Rosenstock et al., 1988), and more recently technology (Chaudhry et al., 2006): the Internet is significantly contributing to patient experiences (Powell et al., 2003; Rice, 2000; Eysenbach et al., 2004), as reflected in NHS England's "Widening Digital Participation" programme.

However, the growing medical boundaries also need to transcend to the realm of making data available for research. The HES data application for this research was originally rejected by the HSCIC on the grounds that it would not have a medical purpose (HSCIC, 2014). Though not generating (bio-)medical findings, this research has produced insights of relevance to the health system. The view on what is medically relevant should therefore be extended across all arm's-length bodies of the Department of Health and other stakeholders in the health system, particularly for the further reuse of secondary data, as in SM research.

In addition, data should be made more openly available for research without time-consuming access procedures, which limit research planning and prevent exploratory data analysis at the beginning of the project (Connelly et al., 2016). All census data in this research (Census, 2011a,b,c,d,f,g,h) is openly available down to the OA level, including sensitive items such as long-term conditions and self-rated health. The data is protected through aggregation on the OA level, the collapsing of categories, small number suppression, and record-swapping to disguise unusual individuals and households (ONS, 2011b). In contrast, HES data – at the same level of aggregation – was protected through a time-consuming access process evaluating the purpose of the analysis and security measures (see Appendix XII, p. 339). Naturally, health-related procedural data may deserve more protection as patients may not be aware of the uses to which the data is being put (Brown et al., 2010; Lowrance, 2003), and making processes difficult also ensures that only those with a genuine interest in the data actually get access. Secondary uses of medical data have been discussed extensively in the literature (Hoffman, 2010; Di Iorio et al., 2014), as, particularly for individual-level data, several techniques exist for de-identifying individuals (El Emam et al., 2015), which may decrease trust in the health system and have adverse health effects (Whetten et al., 2006). However, *aggregated, recoded data with swapped records* should be made more openly available for scientific projects, thereby sensibly balancing individual privacy with the advancement of research.

Finally, other datasets should be released for research purposes. Hospital inpatient data usually relates to severe conditions, and suffers from the limitation that the GP acts as a gatekeeper (see section 1.2). In that sense, it is only an imperfect

surrogate for health service use. However, several interviewees referred to recent stays in hospital, so that the data is still applicable for the purposes of this research. Ideally, however, datasets about GP visits should be made available as well. While CPRD exists as a resource in England, it covers only part of the GP practices, and is substantially more expensive to obtain than HES data. This again links to making data more openly available, especially given that the UK is at the forefront of the open data movement (Davies, 2013) and aims to assess the influence of technology on all kinds of health service use (Tinder Foundation, 2013).

7.5 Methods of enquiry within the pragmatic paradigm

The mixed methods design provides an example in which while "methods have been integrated, the whole can be greater than the sum of the parts" (Barbour, 1999, p. 40). The quantitative data enabled viewing the larger picture of Internet use and health, although it may be "weak in understanding the context or setting in which people talk" (Creswell and Clark, 2011, p. 12). This was offset by qualitative research, which is sometimes accused of providing only anecdotal evidence, and lacking generalisability (Mays and Pope, 1995). The mixed methods design therefore made it possible to approach the RQs from different directions, pragmatically taking advantage of the best available methods (Johnson and Onwuegbuzie, 2004).

Then again, the use of simulations and secondary data, the patient-centric perspective of the qualitative research, and the way by which participants were selected in the qualitative part also entailed some limitations, which will be discussed as part of the mixed methods design (subsection 7.5.1), and the qualitative and quantitative data (subsections 7.5.2 and 7.5.3), and will provide the foundation for recommended additions to the secondary datasets in subsection 7.5.4.

7.5.1 Mixed methods research design

Beyond the above-mentioned benefits of the research design, the combination of qualitative and quantitative approaches has proven especially fruitful with secondary simulated data. Despite the benefits of reusing surveys, such as obtaining large-scale insights without tedious data collection (Freese, 2009), "surveys [are] often frustrating to work with given how little we can know about the contours of an individual case" (p. 258). Following up the same individuals quantitatively added a unique analytical perspective to the quantitative data (Pope and Mays, 2006), such as the qualitatively inspired typology, the interaction effects, and the SEM models

developed in this research. In addition, while research usually faces a trade-off between representativeness and saturation due to finite researcher resources (Teddlie and Yu, 2007) particularly when triangulating multiple methods (Patton, 1999), using qualitatively enriched secondary data made it possible to spend more time on the qualitative research. Finally, particularly as the final quantitative dataset was simulated (as opposed to more established research methods in the social sciences), the qualitative data also helped to substantiate the quantitative claims.

The mixed methods design also followed from the intention to juxtapose qualitative insights to understand the mechanisms, and generalisable data to understand the larger picture of how the Internet impacts health outcomes. Rigorously deriving the research design – and the single methods (see subsections 7.5.2 and 7.5.3) – from the RQs is important to avoid "methodolatry", an excessive preoccupation with methods (Punch, 1998; Janesick, 1994). Hence, for investigating patient-centric outcomes such as perceived health and health service use, patient-centric methods were a logical consequence. At the same time, both outcomes are quantifiable, which is an important conceptual duality to ensure that the methods are well connected with each other and the RQs at hand (Punch, 1998). A limitation of this research is that the quantitative results only partially evaluate the connection between specific actions online and immediate outcomes (which would be observable in experiments, for example), even though these traces were established through the qualitative data in which individuals related cause and effect. In addition, it would have been useful to have longitudinal quantitative data, for example for disentangling age, period and cohort effects – all of which were mentioned by the interviewees, but which could be analysed just as general age-related factors based on the cross-sectional data.

Of particular value was obtaining qualitative and quantitative data from the same individuals. This made it possible to develop the typology both qualitatively and quantitatively, and to evaluate its quality by cross-checking classifications in both. This also revealed its imperfections, with deviations between the qualitative and the quantitative classification for 6 out of 31 (19%) OxIS interviewees. However, it is a common misunderstanding that triangulation serves to "demonstrate that different data sources or inquiry approaches yield essentially the same result", whereas "the point is really to test for such consistency" (Patton, 1999, p. 1193). These resulting divergences provided three insights. First, individuals may transcend the type boundaries over time, in line with the process-based models of perceived health and health service use (see section 2.1). This became evident in Sarah's changed attitude towards the Internet, and Elizabeth's altered view of doctors and the health system

after her stroke (see subsection 4.4.1), and is a limitation that this cross-sectional study cannot fully respond to. Second, the typology lacks specificity due to the non-availability of health-related items such as health motivations and health literacy. Even though health information-seeking is similar to other informational activities online (Blank and Grošelj, 2014), the difference between Norman and Anthony, for example, showed that particular attitudes about health information can distort the picture (see subsection 4.4.1). Third, Susan's and Wanda's different attitudes in the interview provide evidence for the constructed nature of survey and interview data, and highlight the challenges of triangulating them (see subsection 3.4.4).

In addition, maximum variation sampling, the preferred sampling technique when a holistic view of a certain phenomenon is aspired to (Kitto et al., 2008), has greatly benefitted from using the quantitative sample for interviewee recruitment. It enabled the research to take advantage of the diversity achieved through random sampling procedures, and allowed the researcher to know about participant characteristics before the interview (Kemper et al., 2003). Individuals who participated in OxIS in the first place may be different (Babbie, 2009), however, and those sampled from this base for the qualitative part may hence also differ from the general population. Despite this, the interviewees consisted of a diverse group of individuals (see Table 9) in terms of those sociodemographic characteristics deemed analytically important (Sandelowski, 1995b). Additional snowball sampling prevented the research from being constrained by the original OxIS sample.

Finally, the "incentive package" catered to different motivations for participating in the interviews – being compensated, doing something good for charity, contributing to research, and specifically helping with doctoral research – and hence attracted a wide range of individuals. While compensation is not without controversy (Gneezy et al., 2011), several reviews found that participation rates were increased without decreasing data quality (Babbie, 2009; Bentley and Thacker, 2004). In line with Gneezy et al.'s (2011) finding that extrinsic incentives may reduce the intrinsic motivations for producing certain behaviours, some interviewees firmly refused monetary compensation. Interviewees of higher socioeconomic status in particular therefore found the charity option added in the second wave a good compromise.

7.5.2 Obtaining qualitative health data

The interviews enabled various insights into the relationship between Internet use and health, confirming mechanisms and influence factors from the previous literature, and highlighting new concepts around the lifestyle paradox, functions of the

social environment, and types of capital that individuals may build up online.

As described in the methodology chapter and followed up in subsection 7.5.1, the decision to obtain qualitative data in semi-structured face-to-face interviews followed the RQs. Interviews were particularly suitable due to the sensitive nature of health research (Bowling, 2009) and the desire to generalise findings while understanding the mechanisms at work, which is also reflected in the pragmatic paradigm. Face-to-face interviews allowed for better rapport-building, while the semi-structured nature enabled connecting the mixed methods design with flexibility to explore emerging themes (Bryman, 2008; Creswell and Clark, 2011). The circumstance that those who responded to the interview summary *"[could] not think of anything more to add"* (Kathleen, p. 357) indicates that the main points were well captured, although only three individuals responded to the member checks (see Appendix XIX, p. 357).

Nevertheless, the interview setup also had drawbacks. First, while recruiting individuals with a maximum variation sample gave a good overview of the ways by which the Internet affected health, only a few voices could be gathered for each topic: only a few interviewees used online support groups, and only one was fully excluded and had no one to access the Internet for him (Kevin, p. 112). In that sense, saturation for the various health-related Internet uses may not have been reached, even if saturation itself is a problematic concept (O'Reilly and Parker, 2012; Wray et al., 2007, see subsection 3.3.1). Second, the patient-centric perspective restricted the scope to the individual's immediate environment, while the medical professionals' perspective would have added insights around the appropriateness of health service use, and enabled the inclusion of "evaluated health" as another health outcome (Andersen, 1995).

In addition, as several individuals explained that they particularly trusted people from their country (see subsection 6.3.2), the role of a researcher from abroad may have affected their trust. Then again, being a non-British researcher without medical education allowed asking more fundamental questions about individuals' experiences with the health system (Arksey and Knight, 1999; DiCicco-Bloom and Crabtree, 2006), and may have helped to avoid radiating regional or class-related characteristics, which may influence the interview (Manderson et al., 2006; Gray, 2009). Moreover, being a doctoral student (rather than a professor or market researcher) seemed to be perceived as less threatening by the interviewees (Rubin and Rubin, 2004), both in terms of their openness during the interview and for getting access in the first place (see subsection 3.3.1). In addition, being a relatively young female researcher may have shaped the interviewees' positioning of themselves, for example

Matthew referring to himself as an old person at the age of 47 (p. 194), several older females presenting their Internet skills as insufficient (Hargittai and Shafer, 2006), or females voluntarily adding more information about their health-related complaints in the interview – a behaviour that is commonly observed for female participants talking to female interviewers (Padfield and Procter, 1996).

This links to the wider discussion about how both the qualitative and the quantitative data may be socially constructed (Creswell and Miller, 2000; Bryman, 2008, see subsection 3.4.4). In general, the interviewees were very open, talking about experiences they had kept to themselves for decades (Ian, p. 151) or online behaviour they deemed personal (Helen, p. 169; Michael, p. 175). Their accounts also highlighted differences that may be traced back to certain role sets in their lives. For example, some older women particularly highlighted their low skills (Shirley: *"I'm not wise up enough to know how to do things"*, p. 120; Karen: *"There's people that are brighter than I am [...] I don't want to wipe [my husband's] stuff off"*, p. 121). While this may partially have corresponded to the interviewees' actual skills, other research found that men and women do not differ greatly in their actual online skills, although women's self-assessments are significantly lower (Hargittai and Shafer, 2006; van Deursen et al., 2011). Perhaps this also explains why those types with lower skills – delegators and adigitals – are composed of more females, whereas all other types have a neutral gender balance (see Table 13).

Finally, several interviewees had stories about other people – family members, friends, or distant acquaintances – who relied heavily on online health information, which reflects the fact that most people know at least one hypochondriac, but usually do not admit to being one themselves (Dillon, 2009). This may indicate that some interviewees may not have fully admitted the effects of reading health information online, for example. This was mitigated by asking various questions and thereby approaching the interview topic from several angles (Krumpal, 2013), and asking for specific instances of health information-seeking, health service use and changes to health perceptions, rather than general attitudes (see interview guide in Appendix XVII, p. 349). In addition, the quantitative data was collected without the specific focus on the RQs in this research, so that the joint examination of both data sources may further offset this effect.

7.5.3 Spatial microsimulation for combining datasets

SM enabled quantitative analyses without the need to collect a large-scale random sample. Using this technique was necessary as existing surveys did not cover Internet

use and healthcare in sufficient detail: OxIS, with rich information on Internet use, contains no health outcomes, while the UKHLS, the OLS and the AMU survey, with more items on health and wellbeing, measure Internet use only with a single item.

As the quantitative results showed, SM is a way to understand relationships on a larger scale. While internal and external validation (subsections 3.2.3 and 3.2.4) confirmed the general suitability of the simulated dataset, Chapters 4–6 evidenced that sociodemographic variables behaved as one would expect from existing research (see section 2.4). At the same time, the qualitative and quantitative results largely converged, as discussed throughout this chapter. While, from the perspective of the pragmatic paradigm, none of the strands may be seen as "the truth", the qualitative data presents a more traditional and widely approved data collection method, and hence may add to the validity and reliability of the SM data.

However, simulated data entails some side effects which complicate its analysis and interpretation. First, this research did not employ significance values, as due to the large-scale simulated data, small and not necessarily meaningful effect sizes would have been significant (see subsection 3.4.3). Even if p-values also face limitations (Carver, 1993; Kirk, 1996; Anderson et al., 2001), and scientific conclusions and policy decisions should not be based on reaching specific thresholds (Wasserstein and Lazar, 2016), not having a diagnostic boundary for when relationships "count" means that only the effect sizes of Internet use can be compared to various sociodemographic variables. This is acceptable as effect sizes are the most important result of scientific inquiry (Cohen, 1990), although ideally, p-values should at least be presented as an additional piece of information (Carver, 1993; Sullivan and Feinn, 2012). Similarly, statistical measures such as the Adjusted R^2 were more difficult to interpret, as they mostly explained variance stemming from how the outcome variables were formed based on the sociodemographic constraints. It is still interesting that Internet use explains part of the variance (even if this was only 1 or 2 percentage points, see Appendix XI, p. 334).

Extending this point, Appendix XI (p. 334) juxtaposes all regressions in this research for the simulated data and OxIS data with directly appended outcome values for perceived health and health service use. While this evidences that only some coefficients were significant based on the OxIS sample, it also showed that the coefficients obtained do not alter the main conclusions derived from the data. To a certain extent, this is not surprising, as SM does not create any new data (Lovelace and Dumont, 2015), but this check contributes to the validity of this research, confirms the applicability of SM for combining secondary datasets (Deetjen

and Powell, 2016a), and thereby presents another methodological contribution of this thesis. However, SM is still superior, as it enables randomising the error (see subsection 3.4.3 and Appendix XI, p. 332) and increases the applicability across England, as evidenced in the case study on the Isle of Wight in subsection 6.3.3. At the same time, the case study of this particular local authority district confirmed that a synthetic population derived via SM preserves the essential structure of the relationships between Internet use and health outcomes, independently of whether the data was integerised as part of the SM process or not.

It is worth noting that SM is a geographical technique of small area estimation (Haslett et al., 2010; Rahman et al., 2010), reflecting the importance of place (see subsection 2.4.3). The ideas behind geographic and statistical models are similar, and their use is largely driven by disciplinary backgrounds (Whitworth, 2013). However, SM has two advantages: it estimates full individuals instead of only single variables, which is useful when dealing with individual and contextual factors, as in this research; and it can incorporate cross-tabulated constraints, for example age by education (Tanton and Edwards, 2013). In addition, IPF as one of the possible SM approaches has the advantages of being simple, fast, relatively accurate, while it also avoids local optima in the simulation (Lovelace and Ballas, 2013).

Finally, while the maps in section 6.3 provide insights into area-based differences for the relationship between Internet use and health, they suffer from the MAUP, and may only give indications about the area-based distribution of the regression coefficients presented. Some area-based patterns may have been influenced by which OxIS individuals were selected, and reinforced by including locality as a constraint in SM (see Appendix V, p. 305). Then again, assuming that the sample is truly random, it would be correct that certain individual characteristics are stronger in some regions; it is simply that the borders between regions, for example, may look artificially sharp based on the way that the simulated data was created.

Beyond the Internet and healthcare, there is an increasing tendency to provide both surveys and general area-based data with location information[12], as evidenced on data.gov.uk or the UK Data Service. The larger availability of geo-located data gives rise to SM and related techniques across disciplines. While location information is often only available with specific permission, SM may be seen as a privacy-preserving method of generating data, as individuals in the final dataset are

[12] For example, the Community Life and Life Opportunities surveys, and the Audit of Political Engagement (ONS), but also broadband speeds from Ofcom, or crime statistics

simulated. In the UK, the ONS considers simulation techniques for creating more frequent updates of the census, or potentially replacing the census altogether (UK Government, 2002). With more data being generated in academic research and from routine sources within and outside health systems, SM may gain in importance as a privacy-considerate and cost-effective way of doing research.

7.5.4 Recommendations for secondary data

While secondary data used in this research is to only a limited extent modifiable or extendable by its nature, the findings have led to recommendations about additions for future research waves, bearing in mind that additions cost resources, add participant burden, and may lower response rates (Dillman et al., 1993). These suggestions may be more immediately addressable for OxIS (if it is continued beyond 2013), but less so for the census and HES data, particularly as the latter is a routinely collected by-product.

For **OxIS**, this research suggests four extensions. First, while OxIS by definition focuses on Internet use, more questions on health would be useful for health-related analyses, and for improving links with other purposively and routinely collected health datasets. Although OxIS asks about health-related Internet use (Q.C21g, Q.C22d), the existence of a long-term health problem, and its effects on Internet use (Q.D16, Q.D17), it would be useful to provide these on a Likert scale (rather than just binary one) and possibly include information on the kind of health conditions (for example, mental versus physical health, based on section 4.1).

Secondly, in the spirit of the outcomes divide, it would be useful to have more items on effects of use. While OxIS has 36 items on uses, it contains only 9 items on the effects of use. For example, "finding information that helped you improve your health" (Q.C22d) could be complemented by a second item focusing on health service use, particularly with the "digital by default" vision of the UK. At the minimum, existing questions around the effects of use should be extended to all respondents. Classifying individuals into separate categories of user, ex-user and non-user at the beginning of the survey interview is not practical for capturing their diverse digitally mediated environment. Extending these outcomes to non-users would enable further research on proxy use and delegation, and facilitate the inclusion of Internet-based outcomes not related to actual use.

Third, some existing constructs should be improved in future years. It would be useful to have a proper self-efficacy scale – currently there are four items (SC.5d–SC.5g) that do not hold together as a construct. This would be a worthwhile addition

for other applications of Bandura's (1986) social cognitive theory across disciplines (for example Eastin and LaRose, 2000; LaRose and Eastin, 2004; Hoffmann et al., 2015). Moreover, additional questions on health behaviour and health literacy would enable further valuable analysis. For this research, these additions would help to confirm the established typology with more health-related constructs, thereby reducing the classification error described in section 4.4.

The fourth and final suggestion for OxIS can be implemented relatively easily: constructs in OxIS should correspond to the ONS constructs as closely as possible. While the construct for NS-SeC that has been developed in this research has already been implemented in the standard OxIS dataset according to the specification in Appendix V (p. 304) from 2011 onwards, this suggestion also applies to age groups and the educational qualifications. Incorporating the definitions corresponding to the census items in the main OxIS dataset would not only facilitate combining the data, but also ensure a consistent means of doing so.

While the **census** has two good health-related questions, used in this research, it would be useful to include a question on Internet use. A use/non-use question would make it possible to use SM with online surveys, as using the census data for users only would then match the survey's sampling frame. However, based on the shifting digital divide (van Deursen and van Dijk, 2014; van Deursen and Helsper, 2015), it might be useful to ask further questions about digital uses and the dimensions of the digital divide around interest, skills, usage opportunities, devices and experience (van Dijk and Hacker, 2003). This would help to corroborate existing work about the geographic distribution of digital inequalities in the UK (Go ON UK, 2015) and support a shift to the "digital by default" vision in the UK. As the future census may be based on simulations and predictions of existing census data and additional surveys (UK Government, 2002), which adds an additional layer of complexity that has so far not been considered in the research literature, these additional data items would also help to evaluate the resulting dataset.

From a broader perspective, it would be useful to standardise the census beyond England and Wales: asking the same items in the Scottish and potentially Northern Irish censuses, and harmonising the definition of OAs across the UK. The same applies to HES data, which in its current form is available only for England. This is not a question of quantity, but would allow a breakdown of the results for other geographies, and include features of the population or health systems in comparative analyses. However, this may create organisational complexities beyond data and data definitions, making the feasibility of this recommendation low.

Due to **HES data** being routinely collected, it would be against its nature to add further items beyond hospital operations and management. However, in addition to the general points about making HES data available (see subsection 7.4.3), it should be made easier to obtain HES data (such as inpatient/outpatient visits, diagnoses) split not only by age and gender as in this research, but according to the census categories for education, NS-SeC and long-term conditions. Individual HES data is already linked to individual census data – but on the grounds of privacy protection, the hurdles for linked data are even higher than is currently the case (see subsection 7.4.3 and Appendix XII, p. 339). However, this would give rise to further SM research: either via the census with the key outcome variables being better broken down by the constraints (see Appendix VI, p. 306), or with HES data being used directly for SM (which was not possible in this research due to a lack of constraint variables in HES data – see Table 4 and Figure 7).

However, without these additional items, the secondary datasets employed have provided insights into the relationship between Internet use and health, particularly through the qualitative enhancement which also inspired many of the outlined suggestions. The main findings for each of the RQs, alongside the most important contributions and ideas for further research, are synthesised in the next and final chapter of this research.

8 Conclusion

Internet use does influence how individuals make use of health services, and how they perceive their health. However, to what extent one influences the other, and what this influence looks like, varies across individuals depending on their previous motivations and enablers (Chapter 4), situational factors (Chapter 5), and individual and contextual circumstances (Chapter 6). In addition, the links by which Internet use, perceived health and health service use are related are complex: there are various mechanisms, apart from their mutually causal links in general.

Consequently, being online does not automatically come with health benefits, nor does being offline equate to being unhealthy. While the online realm creates ways to obtain official or user-written health information, helps embeddedness with friends and family, and enables the forming of new contacts, not everyone makes use of this – and the effects vary across individuals. Particularly for those who cannot use the Internet independently, the social environment plays an important role for filtering information and compensating for their lack of enablers. At the same time, some mechanisms through which the Internet influences perceived health and health service use also exist offline, so that the Internet – particularly for the content and connection affordances – may in fact represent new ways of doing old things from the perspective of health outcomes. Yet, other Internet-based health outcomes exist for the mediation affordance of the Internet, and thereby relate to actual use.

In the context of England, these insights have implications for policymakers and practitioners for the ongoing transition to a "digital by default" vision while leaving nobody behind. The circumstance that these insights are partially derived from

secondary data combined through SM also has implications for how data providers make data available. These contributions serve as a launching point for further inquiry: in addition to other methodological approaches and longitudinal analyses for replicating the results, other research should build the established typology with health-specific dimensions and analyse its applicability to further Internet outcomes, follow up emerging themes – such as the lifestyle paradox and the outcomes divide – and incorporate the complementary physicians' perspective.

This final chapter synthesises the results for RQ_1-RQ_3 in section 8.1. Section 8.2 summarises the contributions of this research to the existing theoretical work, its practical value and its methodological insights, before directions for further research in the realm of Internet use and health are laid out in section 8.3.

8.1 Synthesis of research questions

To separate out the different aspects of how the Internet influenced perceived health and health service use, RQ_1 asked to what extent Internet use, and particularly health information-seeking, influenced these two outcomes for both users and non-users, whereas RQ_2 aimed at disentangling its underlying mechanisms. RQ_3 served to understand how individual and contextual factors shape Internet use and health, as well as their relationship. This section synthesises the key findings for each RQ.

8.1.1 RQ_1: Health information for (non-)users

Being an Internet user is generally associated with higher levels of perceived health and lower levels of health service use. This finding holds true after controlling for age, gender, education, socioeconomic status and long-term conditions. The qualitative and quantitative data confirmed that this link is bidirectional: while health-related factors prevent some individuals from using the Internet, and health problems and health service use may induce health information-seeking, being an Internet user also influences how individuals perceive their health and make use of health services, which is the main purpose of this research, and the focus of RQ_1.

However, not all users and non-users are equal, and the lines between both groups are blurry when it comes to health outcomes. For health information, this research showed that outcomes vary particularly across different types of individuals – termed the learners, the pragmatists, the sceptics, the worriers, the delegators and the adigitals in this research. This classification rests upon prior motivations for health-related Internet use, including individuals' learning attitudes, attitudes

towards the Internet's usefulness, enjoyment of being online, and trust in doctors, and enablers such as skills, technology attitude, Internet interest and self-efficacy.

The learners wanted to find out as much as possible about their and others' health, and thereby reduce their health service use, shape their consultations, and improve their health perception. The pragmatists primarily used the Internet as a quick screening tool, thereby particularly influencing their health service use, although these feelings of control may also affect their health perception. In contrast, the medical sceptics embraced the Internet to avoid health service use, and while the data confirms that this may be the case, they may also have lower levels of perceived health. For the worriers, as the group with the highest frequency of health information-seeking, the Internet led to more concerns. While there was barely any evidence of increased health service use, the findings indicated that the worriers did at least not reduce their health service use based on health information online. The delegators preferred to have other people find health information for them, and while these social filtering mechanisms led to lower concerns, health information online had nearly no impact on their frequency of health service use. For the adigitals, there was essentially no effect on either health outcome.

While the first four types are largely made up of users, and the latter two of non-users, some individuals transcend these boundaries. Those non-users in the first four types are primarily held back by the lack of a suitable device or by low skills, but in terms of their motivations rather correspond to interested users. In contrast, those users in the latter two types often also lack these enablers, but primarily are not interested in using the Internet in general or for health-related purposes – with the difference that the delegators generally see the value of the Internet, as opposed to the adigitals. As a consequence, the relationship between Internet use and perceived health was as strong for users of these latter two types as for non-users. Particularly for the delegators, the relationship with perceived health was equally strong for users and non-users, which may point to the delegation aspect that interviewees of this type highlighted.

Consequently, the typology of health information-seekers is useful for understanding the Internet's effects on perceived health and health service use. It illustrates that the digital divide is increasingly more complex to delineate (van Dijk, 2005), shows the role of antecedent factors such as motivations and enablers (van Deursen and van Dijk, 2014; Helsper, 2012), indicates that previously existing health behaviours also transcend to the online realm (Hale, 2013), and suggests an "outcomes divide" in terms of the benefits that individuals obtain online (van Deursen

and Helsper, 2015). However, the typology is a tool to understand orientations towards Internet use and health, rather than a rigid structure which unequivocally places individuals, who exist in a complex social reality, in boxes. Individuals may fall into different types based on specific life events or changed attitudes over time, may exhibit characteristics of more than one type in certain situations, and may have specific predispositions about health information not fully captured by the typology formed based on general motivations and enablers.

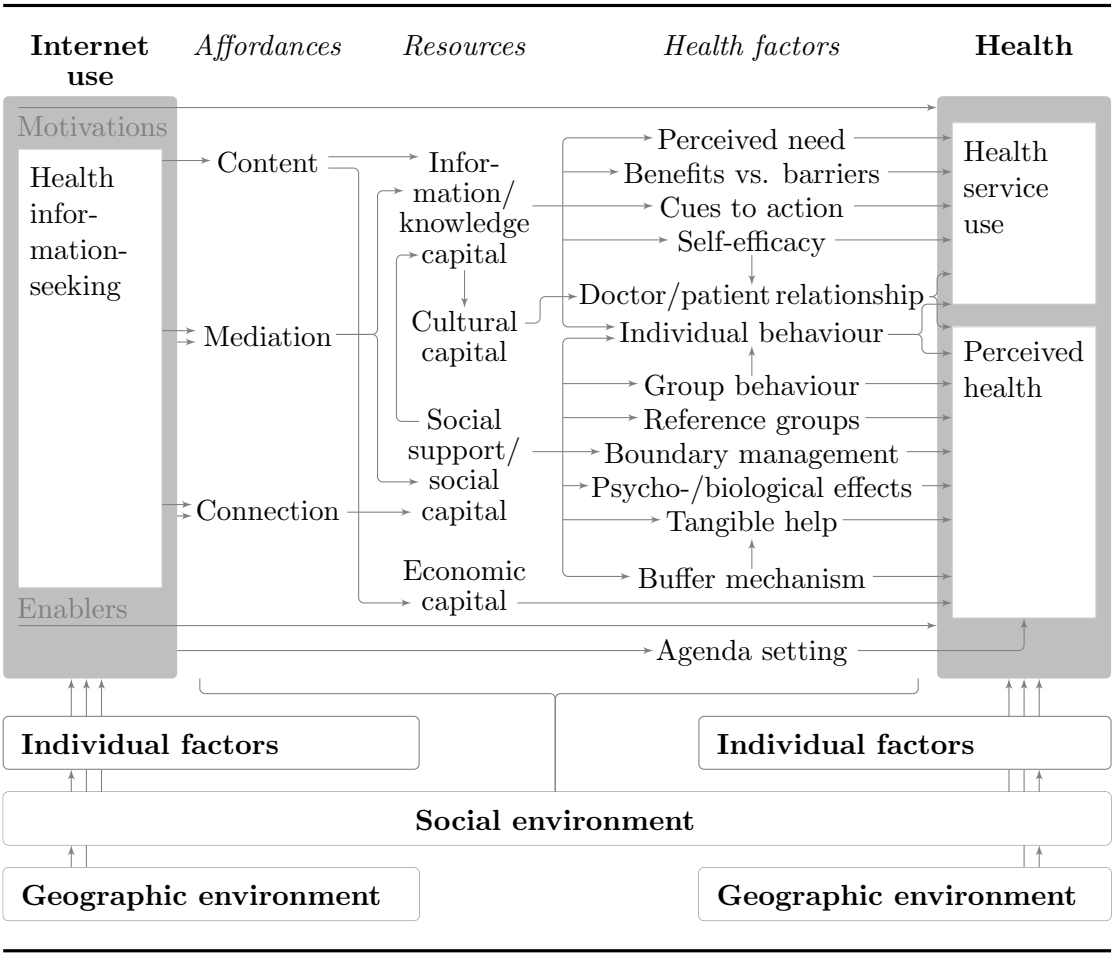
8.1.2 RQ_2: Mechanisms of Internet use and health

Health service use was primarily influenced by health information online, which represents the Internet's content affordance. The Internet reduced health service use when the symptoms were clearly nameable and not particularly severe, helped to evaluate urgency, and enabled the monitoring of the elderly or those with long-term conditions to avoid unnecessary health services, but seek help quickly when needed. Then again, health information online also encouraged health service use for stigmatised conditions, as people learned keywords online and felt legitimised to seek help, while there was only anecdotal evidence of unnecessarily increased health service use. Moreover, health information online shaped the consultation itself: it spurred discussions in the consultation, enabled individuals to be more demanding if required, and enabled more informed choice for those considering different treatment options or deciding to have certain procedures. On a larger scale, the Internet made some individuals aware of the positive aspects of the current health system, and thereby improved their satisfaction with the NHS in general.

For **health perception**, the Internet helped to re-evaluate situations, mostly with positive effects, except when users came across exaggerated or incidental findings. The Internet also enabled comparisons with others, and thereby contributed to feeling normal or part of a bigger group. Some also highlighted the benefit of user-generated health information, which affected their health based on changed behaviours. However, Internet use also created a sense of guilt due to a constant awareness of lifestyle topics, thereby shifting what was considered "perfect health". In addition, the Internet helped individuals to manage boundaries between online and offline identities, thereby creating spaces to be free of certain diseases. With positive effects on health, the Internet also enabled the finding of companionship and embeddedness with family and friends, although it may have further isolated some due to creating addictive habits. Finally, the Internet may change general life circumstances, positively affecting self-perception and perceived health.

These mechanisms confirmed that all of the Internet's affordances (content, communication and mediation) are relevant in the health context. Through these affordances, individuals build up resources such as knowledge capital and social capital, even though the term "capital" in the sense of long-lasting resources could not always be sustained here. This research also found resources in addition to those derived from the literature: cultural capital (through learning keywords and facts for the consultation) and economic capital (through learning about benefits, or finding jobs due to knowledge obtained online). Beyond factors influencing health service use and perceived health based on previous research (see section 2.1), this research identified agenda-setting, reference groups, and boundary management as important additions to the analytical framework. Although the doctor-patient relationship was outside the scope of this research, the findings confirmed its central role for understanding the effects on health service use and perceived health. Figure 30 shows the updated analytical framework.

FIGURE 30: Updated mediating mechanisms in analytical framework



Another finding for RQ_2 is that some mechanisms also exist offline: content may be obtained from other sources or shared offline, and offline connections enable the exchanging of experiences and emotional support. In that sense, only those mechanisms stemming from the mediation affordance – those that rely on personal, unmonitored access, or the availability of information at different places – are unique to the Internet, so that to some extent, the Internet may be a different way to fulfil pre-existing needs (Gershuny, 1983).

Finally, the "lifestyle paradox" that this research uncovered is that Internet use may decrease health perception due to an awareness and guilt with respect to fitness and lifestyles promoted online. Consequently, Internet use may generally be positively related to perceived health, as described in the context of RQ_1, but not necessarily to the highest level or "excellent health".

8.1.3 RQ_3: Individual and contextual influences

The analyses for RQ_3 showed that individual factors are largely related to antecedent factors around motivations and enablers, and primarily acted as moderators. Older individuals often exhibited lower Internet interest, and found the Internet less important for information – although for those interested against the odds, Internet use was related to higher perceived health and lower health service use. The more educated displayed more trust towards the medical profession, and better skills for finding and evaluating information. Lower socioeconomic status was often associated with lower technology attitudes and lower self-efficacy in terms of contributing to one's own health. Self-efficacy was also what differentiated those with long-term conditions, and was generally positively related to perceived health. While the typology of health information-seekers showed differences in socioeconomic status, individual factors did not significantly predict type assignment. An implication is that outcomes inequalities do not necessarily follow the known sociodemographic facets of the digital divide (van Deursen and van Dijk, 2014).

While the social environment affected these individual factors, as well as the associated motivations and Bourdieu's (1986) habitus, it also shaped the relationship between Internet use and health. First, it acted as a support mechanism to offset a lack of enablers, both in terms of using the technology, but also in finding medical keywords to use with a search engine, particularly for those with low information-evaluation skills. Second, the social environment also represented a filter to remove potentially frightening information in the current situation. Finally, the Internet also enhanced social embeddedness for some due to maintaining or establishing

new connections online, with positive effects on their health perception. This also demonstrates the dual role of the social environment as a moderator and mediator for the relationship of Internet use and health outcomes.

Finally, the geographic environment – similarly to the individual factors – again influences Internet use, perceived health and health service use separately, rather than their relationship. One exception may be that the Internet's ability to reduce loneliness was more strongly associated with perceived health in urban areas, where social exclusion is more a problem (Power, 2000). In addition, many individuals highlighted the importance of being in the UK, using UK resources, and talking to other UK citizens about their health concerns, illustrating how the Internet may be perceived locally despite its global nature (Hampton and Wellman, 2003).

The influence of individual and contextual findings has also been updated in the analytical framework in Figure 30. With the finding that only the social environment really influences the relationship between Internet use, perceived health and health service use, individual factors and the geographic context are still important to incorporate due to their influence on each of the concepts separately. Additionally, these findings support the choice of a geographical approach to combining different data sources through SM, and confirm the selection of constraint variables.

8.2 Contribution of findings

Analysing how the Internet influences these two health outcomes helped to advance the theoretical understanding of Internet-based outcomes, derive recommendations for practice and policy, and apply SM to connect secondary datasets.

8.2.1 Theoretical contributions

The first contribution of this research concerns the mechanisms by which the Internet influences perceived health and health service use. This includes the insights that all three affordances matter, that beyond information and social capital also cultural and economic capital play a role, and that in most cases these may not be "capitals" in the sense of a long-lasting resource – except for learners and some of those with long-term conditions (in terms of information capital) and those with existing social networks that they maintain and develop online (in terms of social capital). In addition to the paths identified using previous models in the literature review, this research added agenda-setting, reference groups, boundary management and the doctor/patient relationship as additional influencing factors.

The second theoretical contribution is the finding that segmenting individuals by their motivations and enablers for Internet use helps to redefine the user/non-user split in the context of health outcomes. A corollary is that the Internet reinforces prior (offline) health behaviours, rather than introducing something entirely new. Particularly those mechanisms enabled by the content and connection affordances also exist offline, whereas the mediation affordance is particularly novel: enabling private and undisturbed access (for embarrassing health questions), cue reduction (for open conversations about health issues), and mass communication and crowdsourcing (for efficiently consulting one's network for health advice), as well as enhanced modalities and increased geographical reach (for being in contact with family and friends, or watching documentaries on health systems).

The third and related contribution is the shift to outcome inequalities. While inequalities in terms of access, support and usage divides (DiMaggio and Hargittai, 2001; van Deursen and van Dijk, 2014; Warschauer, 2004) remain relevant, additional inequalities exists in terms of the benefits that individuals reap from the technology (van Deursen and Helsper, 2015). This research has proposed an outcomes divide following the typology for health information-seekers, and the health outcomes they obtained online based on their motivations and enablers. As opposed to the notion that access divides diminish with growing Internet penetration (DiMaggio and Hargittai, 2001), this research showed that some gap persists for those non-users who form part of the learners, pragmatists, worriers or sceptics groups, and primarily do not use the Internet due to a lack of skills or devices. In addition, while proxy use is beneficial in the health context, those without access may not obtain the benefits that mediation enables, and those without sufficient support may obtain reduced benefits due to difficulties finding and interpreting health information.

The fourth contribution relates to the interplay of agency and structure. This research showed how structure – in the sense of sociodemographic characteristics – shapes agency in one's own health, but also how agency shapes structure, for example through improving life circumstances with positive effects on health. Moreover, this research illustrated how the field of the professional consultation is redefined through the Internet, while the Internet also enabled the management of different private and professional fields. For users, the Internet makes spatial structure of less importance for who they compare themselves to, although their geographic environment remains important in terms of its contextual health-related and compositional features.

Finally, a fifth contribution of this research is the lifestyle paradox. While Albrecht and Devlieger's (1999) disability paradox describes how individuals may

perceive their health as very good despite having health issues, the lifestyle paradox describes the reverse: individuals rate their health as lower as the Internet creates a sense of lifestyle awareness and potentially even guilt, which raises the bar for what "perfect health" means. This paradox is an important contribution to agenda-setting theory and media theory more generally, as it shows an unexpected effect. This effect may be seen as negative due to lower health perceptions, but may also have positive influences driving individuals towards more healthy behaviour, which, as explained in the literature review, is an important factor influencing health.

8.2.2 Practical contributions

For **policymakers**, the implications of this research are three-fold. First of all, existing initiatives such as NHS England's programme "Widening Digital Participation" are important milestones to enable Internet-based health benefits and support efficiency in the health system while leaving nobody behind. They address an important skills dimension that became evident in this research, and address those individuals who generally already attribute some usefulness to the Internet. Moreover, some ongoing activities – such as providing certificates of trustworthiness for apps to make them "NHS-approved" – will be highly welcomed by those who form part of the worrier and the pragmatist type, while not creating any disadvantages for the other types who value user-written information as well.

The second implication is around the high importance of social inclusion. While social inclusion itself has positive effects on health (Cacioppo et al., 2002), this research showed the bidirectional importance of social inclusion for offsetting skills gaps as far as possible, and as a social filtering mechanism for obtaining health information. At the same time, the Internet itself may help users to feel more integrated with family and friends and helping them find companionship, although this may primarily be the case for those who attribute positive effects to the Internet.

A third implication is around the introduction of new digital services. While particularly the pragmatists, learners and some sceptics valued the efficiency associated with these services, offering digital services compares to pushing at an open door for these individuals. However, this may not be the case for the adigitals. Understanding these differences and tailoring interventions accordingly, and addressing all enablers, including interest and usage opportunities (rather than focusing on skills and devices only) is paramount. At the same time – and as is done in the "Widening Digital Participation" project – addressing skills and devices is particularly important for those who would be interested in using the Internet.

For **practitioners**, the main implications are that they need to actively embrace the benefits that the Internet may offer. While individuals are different – as argued in the context of the typology developed here – an overarching implication is that many patients wish to openly talk about their activities online, and use the doctor as a trusted advisor and "sounding board" for information found on the Internet. In addition, some patients would appreciate specific recommendations for online health resources. While the learners and pragmatists navigate the Internet competently and independently, the worriers and the delegators particularly may benefit from trusted advice as a step towards realising Internet-based health outcomes.

At the same time, practitioners themselves may have a role in terms of how new digital services are positioned. While the digital paradigm often highlights convenience, interactivity, increasing relevance, and ease of access for the consumer (Dörner and Edelman, 2015), this research has shown how the adigitals (making up 25% of the population) are particularly critical of new services and equate the shift towards digital services with a deterioration in healthcare. For this, practitioners themselves have to be champions in bringing the technology to patients, and it goes without saying that they need a sufficient level of the enablers discussed in this research (particularly skills and interest) to carry their patients along.

Finally, for **data providers** (such as the HSCIC in the UK), the main implication is that the vast amounts of available data should be made more accessible so that value can be derived from it, for example in aggregate form. While data privacy is important and health data should be protected to avoid endangering individuals' trust in large-scale data resources, data provision guidelines should be broadened in two directions. First, the boundaries of medically relevant data have widened considerably, and go well beyond the confines of the health system (Deetjen et al., 2015). Second, methodologically innovative approaches such as SM should be embraced as novel techniques for privacy-preserving data analysis (Hermes and Poulsen, 2012), rather than being unnecessarily scrutinised on the grounds that they may not belong to standard approaches in epidemiology and public health. Its value as part of the methodological approach will be synthesised in the next subsection.

8.2.3 Methodological contributions

The main methodological contribution of this research is that SM generally works to derive new insights from combining existing datasets, as also confirmed by appending the health outcome values to OxIS directly. While SM has previously been applied in health-related contexts (Anderson and Speed, 2009, 2010; Morrissey et al., 2008,

2010, 2015; Tomintz et al., 2008; Smith et al., 2011), it has primarily been used to estimate the geographic distribution of certain variables, rather than to combine datasets with geographic location as a proxy. Particularly in an era where more data is collected than can be analysed, the benefits of SM as a cost-effective and privacy-preserving method for reusing existing data sources come to the fore. While there are some analytical limitations – such as the non-existence of p-values, the limited meaningfulness of R^2, and the limits of analysing discrete values in the context of the lifestyle paradox – the benefits may outweigh these downsides. This is particularly true for health topics, where data is sensitive and therefore highly protected.

A second methodological contribution lies in the proposed additions to the secondary datasets. For OxIS, additional items on the kind of health problem and health service use would be useful. OxIS should also focus on outcome rather than activity items, and should ask these of all participants due to the blurry distinction between users and non-users. While items such as health literacy, "my doctor" versus "doctors in general" (Hall et al., 2001; Ridd et al., 2009), and health-related self-efficacy may be too specific for an omnibus survey, a good general item on self-efficacy at least should be implemented due to its significance for other applications of social cognitive learning theory to Internet research. Specifically for SM, OxIS constructs should closely correspond to census items. For the census, a question on Internet use might be justifiable given the importance of digital services, could serve as an additional constraint for examining Internet activities and outcomes, and would enable doing SM with online surveys as well. Finally, it would be beneficial to have HES data linked to the census routinely, thereby making it possible to include education and NS-SeC in further SM models.

A third methodological contribution is showing the value of the mixed methods design, as the qualitative insights were indispensable for developing the quantitative models in this research. In particular, using the same individuals in both parts created unusual possibilities for moving between the two data sources. This enabled the typology to be derived both qualitatively and quantitatively with possibilities for cross-checking, and enhanced validity and reliability by substantiating and juxtaposing insights from both methods. Moreover, the qualitative sampling strategy based on the quantitative data allowed for the creation of a maximum variation sample for the interviews. Despite the downside that individuals did not engage in all potential health-related activities online, this ensured a more realistic picture of the "average individual" in England, and decreased the risk of overstating the Internet's significance in individuals' everyday lives.

8.3 Directions for further investigation

While this research has made various contributions, some aspects had to be taken out of the scope, or could not be answered through the chosen methodological approach. Further research should therefore investigate four distinct, but interrelated areas.

First, it would be worth researching the health-related outcomes of Internet use using different methodological approaches. Through including longitudinal data, all three prerequisites of causality (see Hume in Miles and Huberman, 1994) may be validated, and age-related factors may be disentangled into age, period and cohort effects. In addition, while SM presents an innovative way of combining secondary data to derive new insights, it should be verified whether these health outcomes are quantitatively obtained for "real" individuals. This could be enriched, for example, through experimental designs that observe individuals in situations in which they are required make decisions based on health information online. Similarly, focusing on a longer time frame, it should be analysed how the Internet changes individuals' trajectory over the course of specific health conditions, as several of the disease process models suggest (see section 2.1). While this research showed the relevance of the Internet in how individuals experience health and healthcare, this would strengthen the outcome evidence and foster trust in simulations as a resource-effective and privacy-preserving way of developing insights.

Second, as the typology of health information-seekers was formed using only general motivations and enablers in OxIS (see subsections 4.3.3 and 7.5.4), it would be worth replicating it with health-related dimensions: for example, with constructs around health literacy and health-related self-efficacy. Related to the earlier point about longitudinal data, the typology should also be based on measurements from multiple points in time, as individuals may develop and fall into different types following different events in their lives. For the broader context of Internet research, it would be theoretically and practically relevant to understand whether the established typology also describes orientations that are relevant for other Internet-based outcomes. This would not only cross-validate the results of this research, but also help to analyse other Internet outcomes. Finally, with the "cultures of the Internet" established as part of OxIS (Dutton and Blank, 2013), it would be worth researching their overlap for understanding Internet outcomes.

Third, further research should follow up the emergent themes from this research. For example, this relates to the lifestyle paradox in the sense of how the shift to a lifestyle society affects health perceptions, and the bidirectional relationship between

digital and social inclusion as an effect of and prerequisite for Internet-based outcomes. More research beyond the health area should focus on the "outcomes divide", for example following up whether the Internet indeed affects life circumstances more generally. This is an area that deserved more attention as the next frontier of health-related inequalities beyond access, support and usage (van Deursen and Helsper, 2015). Finally, smaller studies should also continue to focus on the health outcomes of specific activities online, for example online support groups.

Fourth, the patient perspective in this research should be complemented by the professional perspective. In line with Hall et al. (2001) and Ridd et al. (2009), this research showed that individuals distinguished between "my doctor" and "doctors in general" (Brian, p. 136) or highlighted qualities of their specific GP (Catherine, p. 122) or GP practice such as waiting times, and showed interest in their doctors' recommendations. Complementing this research through the physician view might therefore be useful in understanding the role of the professional in the process, and enable the inclusion of the appropriateness of health service use (rather than just increase or reduction) or Andersen's (1968) concept of "evaluated health". From a theoretical perspective, this would allow deriving implications for the field of the professional consultation (subsection 7.3.1), refining the analytical framework, and creating further implications for policy and practice.

For all of these exciting and relevant avenues, this thesis serves as a comprehensive conceptual and empirical starting point. By demonstrating that Internet use has an effect on both perceived health and health service use in various ways for different individuals, and substantiating these findings both quantitatively and qualitatively, this research has made an important contribution to the study of Internet-based health outcomes and digital inclusion. Implementing the necessary policy measures, supporting Internet-based benefits through health professionals and within the social environment, and increasing individuals' self-efficacy to participate in their health will help to derive benefits for individuals, the health system and society over the coming years and decades.

Appendices

Literature review

I Overview of previous studies . . . 287
II Key elements of theories on health service use . . . 291
III Key elements of theories on perceived health . . . 294

Methodology

Quantitative

IV Mapping of theoretical constructs to datasets . . . 295
V Constraint variables for spatial microsimulation . . . 302
VI Construction of outcome variables . . . 306
VII Internal validation results . . . 308
VIII External validation results . . . 310
IX Spatial microsimulation in R: approach and code . . . 316
X Regression diagnostics . . . 328
XI Results for OxIS versus simulated dataset . . . 332
XII Timeline of HES data request . . . 339

Qualitative

XIII Area selection . . . 340
XIV Interviewee health conditions . . . 341
XV Overview of interviewees . . . 342
XVI Interview participant documents . . . 346
XVII Interview guide . . . 349
XVIII Coding scheme for data analysis . . . 352
XIX Interviewee responses to research summary . . . 357

Dissemination

XX Policy summary . . . 361
XXI Overview of publications . . . 363

I Overview of previous studies

A structured literature review was used to identify previous studies (Petticrew and Roberts, 2006). This was done using Papers 2.7.3 (a citation management programme), which searches Google Scholar, Google Books, the ACM Digital Library, JSTOR, PubMed and the Web of Science, using keywords such as "Internet use" jointly with "health outcomes", also replacing the former with "online" and "eHealth" and the latter with "health status", "perceived health", "health perceptions", "health service use", "medical appointments", and "physician visits" to cover all concepts. Based on a review of abstracts and follow-up of sources used in the identified papers, the inclusion criteria for papers are the following:

1. Research into the effect of Internet use (including connected software) on any of the following: health outcomes, health status and health service use
2. Published in peer-reviewed journals or as a book chapter
3. Containing empirical research (no reviews), both qualitative/quantitative

This structured literature search was conducted in February 2014. It was updated in March 2016 through a search with the above-mentioned keywords limited to papers published between 2014 and 2016, a forward search on the papers originally identified, and through keeping up to date with publications in several journals in the area of (medical) sociology, psychology and communications.

Outcome concepts	Method	Sample	Limitations	Main findings
Alemi et al. (1996) – pregnancy and drug abuse				
Perceived health Health service use	Quasi-experiment with structured interviews and content analysis of posts	N=53; recruited through medical setting	Use of a voice-bulletin board only, self-reported health service use only, small sample	The experimental group showed a lower use of health services, but at the same time no difference in health status
Berger et al. (2005) – stigmatised conditions (anxiety, depression, herpes, incontinence)				
Health service use	Quantitative; cross-sectional survey analysis	N=7,014; participants from a pre-recruited panel in the US	Focus on stigmatised conditions, self-reported health service use only	Those with stigmatised conditions reported increased health service use based on information and exchange with others online (causal relation)
Bessière et al. (2008) – depression				
Perceived health Depression as a specific health outcome	Quantitative; longitudinal panel survey (3x over 18 months)	N=740; national US sample recruited by random digit dialing	Over-sampled Internet users, hence not generalisable; no control for general Internet uses	Positive effects on depression for communication with friends/family; for communication with strangers, direction of effect based on initial social support level

Outcome concepts	Method	Sample	Limitations	Main findings
Bessière et al. (2010) – depression				
Perceived health Depression as a specific health outcome	Quantitative; longitudinal panel survey (3x over 18 months)	N=740; national US sample recruited by random digit dialing	Over-sampled Internet users, hence not generalisable; no control for general Internet uses	Informational health-related use increases depression, social use reduces depression; no effects of initial health status or primary caregiver status, no change of general health
Casiano et al. (2012) – obesity, depression for young people				
Perceived health Several specific measures (binge drinking, help-seeking behaviour, suicide ideation etc.)	Quantitative; analysis of Canadian Community Health Survey	N=9,137; random sample of only young people aged 15–19	Internet analysed as "en bloc" use without differentiation	Only obesity associated with media use (television/video), no effect for other media uses or health outcomes
Dwyer and Liu (2013) – no condition-specific focus				
Health service use Physician and Emergency Room visits	Quantitative, analysis Community Tracking Study 2000-2001 Household Surveys	N=23,486; multi-stage random stratified sample; telephone and face-to-face interview	No control for health outcomes, therefore no focus on appropriateness of service use	Health info. increases likelihood of physician visit and frequency, but decreases accident and emergency visits; low-trust patients substitute physician services for online health info.
Gustafson et al. (1999) – HIV				
Health service use Utilisation and length of stay, in addition: health-rel. quality of life	Quantitative; randomised controlled trial for 3 and 6 months respectively	N=204; recruited through medical setting	Specific software only; only self-reported health service use; small sample	Positive effects on health-related quality of life and more efficient use of health services (less time during visits, fewer and shorter hospitalisations)
Kraut et al. (1998) – depression				
Perceived health Depression as a specific health outcome	Quantitative; longitudinal survey analysis	N=256; recruited from four school/ neighbourhood groups	Small sample, no control for initial social circle	More Internet use was associated with declines in offline social participation, and increases in depression as a health outcome
Kraut et al. (2002) – depression				
Perceived health Depression as a specific health outcome	Quantitative; longitudinal survey analysis	N=208; recruited from four school/ neighbourhood groups	Small sample	Positive effects of Internet use on depression with a longer follow-up period (3 years) compared to Kraut et al. (1998)

Outcome concepts	Method	Sample	Limitations	Main findings
Kross et al. (2013) – no condition-specific focus, Facebook usage				
Perceived health	Quantitative; longitudinal over 2 weeks through text messages; analysis of lagged effects	N=82; recruitment through flyers in Ann Arbor, Michigan	Very small sample for quantitative analysis, bipolar scale used for perceived health	Facebook use predicted negative shifts in perceived health and life satisfaction more generally (analysing effect at t=2, controlling for t=1 results)
Lee (2008) – no condition-specific focus				
Health service use	Quantitative; longitudinal study (over 1 year)	N=4,957 in first wave, N=1,486 in follow-up wave	Access barriers (health insurance) were not controlled for in this US study; low response rate at t=1	Online health information at t=1 increased the frequency of contact with health professionals for information and for ailment treatment at t=2, while controlling for current health status
Nabi et al. (2013) – Facebook usage				
Perceived health (both physical and mental health)	Quantitative; cross-sectional survey	N=401; recruited in university setting	Skewed student sample, specific case of Facebook only	Social use (Facebook) led to less physical illness / greater well-being; no association with intensity of use, but number of friends online, weaker when adjusted for offline social networks
Rains and Keating (2011) – focus on bloggers				
Perceived health	Quantitative; survey to research social support effects on perceived health	N=121 health bloggers; identified online via search	Support from blog readers and family/friends enmeshed, different conditions have different support needs	Health blogging presents an opportunity without existing support to obtain social support online, which in turn had positive effects on their health perception
Rice (2006)				
Perceived health Health service use	Quantitative; cross-sectional analysis of secondary survey data	7 surveys of different sample sizes, including Pew Internet survey and American Life project	Only use of patient-rated health and patient-reported health service use	Positive influence of health information-seeking on perceived health, overall low influence on health service use. Best predictors for health outcomes: gender, more frequent health information seeking, and medical characteristics

Outcome concepts	Method	Sample	Limitations	Main findings
Suziedelyte (2012) – no condition-specific focus				
Health service use	Quantitative; cross-sectional analysis of secondary survey data	N=16,677; data from US Health Information National Trends Survey	Self-reported and very wide definition of health service use in a single variable	Higher health service use for those who seek information online (after controlling for age, gender, race, education and marital status)
Whitacre and Brooks (2014) – focus on broadband adoption				
Health service use Various additional outcomes: Flu shots, drinking, asthma, diabetes, exercise etc.	Quantitative; area data from Center for Disease Control and Federal Communications Commission	N=92 areas, not individuals, all areas from the original dataset, excluding those with missing data	Area-based data only, mechanisms between broadband (in contrast to Internet use) and health not entirely clear	Broadband adoption explained (positive/negative) changes in 9/24 outcome measures. For perceived health, broadband adoption was associated with higher likelihood of rating health as "very good" (3/5) instead of "excellent" (4/5)
Ybarra and Suman (2006) – no condition-specific focus				
Health service use	Quantitative; longitudinal data from "Surveying the Digital Future" survey (year 3 and 4)	N=2,104 at t=1; N=1,316 at t=2, all participants $\geq$ 12 years, English/Spanish speaking (US research)	Relatively crude measures of Internet use, no data on reasons/ motivations for actions, high levels of imputation for missing data	As a result of health information, people felt better about information they had received from their doctor, contacted a health professional more frequently, tried to diagnose problems, and sought support from others

II Key elements of theories on health service use

Below is an overview of models discussed in subsection 2.1.1; their relevance or corresponding item in the analytical framework is provided in italics. Bandura's (1986, 2004), which relates both to health service use and perceived health, has not been added separately, as it primarily is a model for health behaviour and is partially contained in Rosenstock et al. (1988).

Author(s)	Name	Description
Andersen (1968, 1995); Andersen and Davidson (2007)	Health service use model	– **Original model (1968):** Health service use is explained by: 1) Predisposing characteristics: demographic, social structure, health beliefs *(=individual, contextual, antecedent factors)*, 2) Enabling resources: family, community *(=contextual factors)* 3) Need: perceived, (evaluated) *(=perceived need)* – **Extended model 1 (1995):** The model was complemented by: 1) Environment: health care system, external environment *(does not apply due to focus on England)* 2) Health behaviour: personal health practices, use of health services *(supports inclusion of health behaviour in model)* 3) Outcomes: perceived health status, evaluated health status, and consumer satisfaction *(supports integration of health service use and perceived health in the same model)* 3) Feedback loop *(supports incorporation of feedback loop in analytical framework)* – **Extended model 2 (2007):** Addition of "Process of medical care" as health behaviour
Rosenstock (1966, 1974); Rosenstock et al. (1988)	Initially no name; later: Health belief model	– **Original model (1966):** Influences on health service use: 1) Perceived susceptibility 2) Perceived seriousness *(1/2 summarised as perceived need)* 3) Perceived benefits of taking action and barriers to taking action *(used as benefits vs. barriers)* 4) Cues to action *(used as cues to action)* – **Extended model 1 (1974):** Extensions by: 1) Demographic variables (age, gender, ethnicity), sociopsychological variables (personality, social class), structural variables (knowledge about the disease, prior contact with the disease, etc.) *(=individual, contextual factors and information)* 2) Perceived susceptibility and perceived seriousness summarised to perceived threat of disease *(supports summarising both to perceived need)* 3) Extended outcome to "Likelihood of taking recommended preventive health action" *(supports inclusion of health behaviour in model)* – **Extended model 2 (1988):** Incorporation of self-efficacy based on Bandura's (1986) social learning theory
Young (1981), as cited in Rebhan (2009)	Choice-making model	– Four factors contribute to health service use: 1) Perceptions of gravity of health problem *(= perceived need)* 2) Knowledge of home treatment *(= benefits vs. barriers)* 3) Faith in treatment *(= benefits vs. barriers)* 4) Access to treatment *(less important in the case of England, but considered as part of individual factors as costs of access may relate to time, psychological barriers etc.)*

Author(s)	Name	Description
Mechanic (1978), as cited in Rebhan (2009)	General theory of help-seeking	– There are 10 main factors influencing health services use: 1) salience of deviant signs and symptoms 2) perception of symptom severity 3) disruption of daily life caused by illness 4) frequency of symptoms and their persistence 5) tolerance of symptoms 6) knowledge/cultural assumptions of illness 7) denial of illness as a result of basic needs 8) whether or not response to the illness disrupts needs 9) alternative interpretations of symptom expression 10) treatment availability and resources, psychological cost *(1,2,3,4,5 = perceived need; 7,8,10 = benefits vs. barriers; 6,9 = information and knowledge)*
Suchman (1965)	Stages of illness and medical care	– The model is primarily a model of the stages of disease with 5 stages, of which particularly Stage 3 is of interest: 1) The symptom experience stage 2) The assumption of the sick role stage 3) The medical care contact stage 4) The dependent-patient role stage 5) The recovery or rehabilitation stage – For stage 3, the following factors influence health service use: 1) Degree to which initial diagnosis / symptoms are accepted *(=perceived need)* 2) Knowledge, availability, and convenience of services *(=benefits vs. barriers, contextual factors)* 3) Social group influences *(=contextual factors)*
Goldsmith et al. (1988)	Help-seeking decision-making model	– The major stages of the help-seeking process are problem recognition, the decision to seek help, and the decision to select specific services. – Major factors influencing the help-seeking process are: 1) Illness profile factors and degree of medical risk: Professional's, individual's and significant other's assessment *(=perceived need)* 2) Predisposing factors: Relative attractiveness of help-seeking, psychosocial resources and usual coping, role oblications: age, gender etc. *(=benefits vs. barriers, individual factors)* 3) Enabling factors: Economic and social facilities, network, accessibility and stigma barriers, social pressures to seek help *(=contextual factors)*
Kleinman (1978)	Medical systems as cultural systems	– The model consists of a professional, folk and popular area. The majority of decisions about health service use (with a professional medical doctor or a folk healer) are made in the popular domain, which comprises individuals, family, the social nexus and community – Decision criteria for health service use in the popular sector are: 1) Beliefs *(antecedent factors)* 2) Roles, relationships, interaction settings, institutions *(individual, contextual factors)* 3) Choices and decisions *(not clearly attributable)*

Author(s)	Name	Description
Dingwall (2001)	Illness as social action model	– Process-oriented, iterative model in which individuals perceive health as a normal state, and return to it by taking action (as opposed to positivist models of straightforward illness behaviour) – Important role of knowledge and experience through which individuals evaluate and interpret their condition *(supports inclusion of knowledge capital and contextual factors in analytical framework)* – Among the possible responses "to the problematic experience can include dismissal (it's not important), 'wait and see' (I'll see how it goes on), self-treatment and consultation with formal health services or others considered knowledgeable" (Wyke et al., 2013, p. 80) *(links to the different outcomes of Internet use, which include health service use as one of the options)* – Outcome is continually evaluated and appraised, also with the help of others – other people may notice a problem, help interpret the problem, contribute their knowledge and experience, and provide tangible help such as buying medicines or making an appointment (Wyke et al., 2013)
Leventhal et al. (2003)	The common sense model of the self-regulation of health and illness	– Iterative model in which individuals have schematic structures about how they respond to health problems, with health service use being one of them; these structures are continually updated through internal and external stimuli that contribute to illness representations – Action taken may be guided by five questions (Wyke et al., 2013): 1) What is the problem (identity)? 2) What is the expected duration (timeline)? *(=perceived benefits of health service use through potential severity)* 3) What causes the problem (cause)? 4) Is it controllable (control)? *(=perceived barriers to health service use by own ability to solve problem)* 5) What are the likely consequences (consequences)? *(=perceived benefits of health service use through potential consequences)* – People are not "isolated problem-solving machines" (p. 54): Self-regulation depends on the input of others *(supports the inclusion of contextual factors in analytical framework)*
Pescosolido (1992)	Health service use rooted in social organisation strategy	– Process-oriented models based on illness careers, as opposed to singular decisions about health service use – High importance of social networks as a way through which individuals learn and understand themselves and shape their help-seeking behaviour *(supports inclusion of contextual factors in analytical framework)* – Shift from individual choice to "socially constructed patterns of decisions" (p. 1096), as opposed to rational choice theory

These three approaches (Dingwall, 2001; Leventhal et al., 2003; Pescosolido, 1992) have also been summarised and compared in Wyke et al. (2013), which helped to create the summary above.

III Key elements of theories on health perception

Below is an overview of the most important meta-studies on influences on perceived health, which provided the foundation to derive the main influence factors discussed in subsection 2.1.2, without any of the numerous primary/empirical studies as this would go beyond the scope of this research. Their relevance or corresponding item in the analytical framework is again provided in italics.

Author(s)	Description
Cohen (1988)	– Review of the major pathways of social support and health, distinguishing between generic models, stress-centered models and psychosocial process models, which again are split up into main effect models and stress-buffering models *(=buffer effects vs. direct effects)* – Within these categories, distinction between four main models: 1) information-based models *(=information factors)* 2) identity and self-esteem models *(=psycho-/biological factors)* 3) social influence models *(=group behaviour)* 4) tangible-resource models *(=tangible help)*
Berkman et al. (2000)	– Review of network effects on health, suggesting four main influences: 1) provision of social support *(=social support)* 2) social influence *(=individual/group behaviour)* 3) on social engagement and attachment *(=individual behaviours)* 4) access to resources and material goods *(=tangible help)* – These network functions influence health through: 1) direct physiological stress responses *(=direct vs. buffer effects)* 2) psychological states and traits including self-esteem, self-efficacy, and security *(=bio-/psychological effects)* 3) health-damaging behaviors (tobacco consumption, high-risk sexual activity) and health-promoting behavior (appropriate health service use, medical adherence, and exercise) *(=group behaviours)* 4) exposure to infectious disease agents (sexually transmitted diseases)
Uchino (2004)	– Main distinction between stress-related models and direct effects models *(=direct vs. buffer effects)* – Stress-centred models (p. 115 in Uchino, 2004): Potential stressor is mediated by received support and available support, which depend on social integration *(=distinction between actual support and available support)*, which in turn affect the potential stressor through esteem, control and coping – Direct effect models (p. 122 in Uchino, 2004): Health is influenced by esteem *(=bio-/psychological effects)* and health behaviours *(=individual/group behaviours)*, which in turn are shaped by social integration, functional support, as well as emotional and social loneliness
Uchino et al. (1996); Uchino (2006)	– Reviews of primarily biological ways through which social support influences health, suggesting that cardiovascular, endocrine, immune system-related changes, and inflammatory processes play a role
Cohen and Wills (1985)	– Focus paper on direct vs. stress-centred models *(=buffer effects vs. direct effects)*, concluding that buffering occurs when social support is defined functionally (assessing "the perceived availability of interpersonal resources that are responsive to the needs elicited by stressful events"), but direct effects occur when it is defined structurally (assessing "a person's degree of integration in a large social network")

IV Mapping of theoretical constructs to datasets

Below are the concepts used throughout the analytical framework (Chapter 2), the overview of datasets (Table 4 in Chapter 3, p. 68) and all analyses (Chapter 4–6) with their quantitative representation in OxIS (2013), HES (2013) and the census (ONS, 2011a). Reversed items are marked with an asterisk (*). For 5-item Likert scales, missing values were recoded as 3 ("neither agree nor disagree").

Internet use

Analytical concept	Quantitative representation	Validity
User concepts: Access divide	Do you personally use the Internet at home/work/ school/college/elsewhere? (Q.H13 in OxIS, 2013), *categorical*	
Next-generation user	– Yes **[AND]** using > 2 Internet applications on a mobile (browse the Internet, use email, update a social networking site, or finding directions) **[OR]** > 2 devices (tablet, reader, multiple computers) as defined by Dutton and Blank (2011)	
First-generation user	– Yes	
Ex-user	– No, but used it in the past	
Never-user	– No, never	
User concepts: Support divide	Have you figured things out by yourself in the last year? (Q.C4g in OxIS, 2013), *dichtomous*	
Independent user	– Yes	
Supported user	– No **[AND]** could probably or definitely find somebody to help (Q.C5 in OxIS, 2013)	
Unsupported user	– No **[AND]** could probably or definitely not find anyone to help (Q.C5 in OxIS, 2013) **[AND]** did not receive any help in the last year from friends/family, at work or in the library (Q.C4a–c in OxIS, 2013)	
	Have you asked someone to send email, get information or purchase a product for you? (Q.N4/Q.E12 in OxIS, 2013)	
Proxy access	– Yes	
Proxy availability	– No **[AND]** knows someone to ask (Q.N6/Q.E14 in OxIS, 2013)	
Unsupported user	– No **[AND]** does not know anyone to ask (Q.N6/Q.E14 in OxIS, 2013)	
User concepts: Usage divide		
Health information seeking	How often do you find information about health or medical care? (Q.C21g in OxIS, 2013), *5-item Likert scale*	
	Note: for "improved my health online" used in the context of the usage divide, see "Internet outcomes" (p. 298)	

Analytical concept	Quantitative representation	Validity
Online activities		
Social activities	Post messages on discussion boards; participate in chat rooms; post photos; read a blog; write a blog; maintain a personal website; Post own writing, stories, poetry or other creative work; make or receive phone calls, do instant messaging; post videos (Q.C8b,c,e-i/Q.C9a,c,d in OxIS, 2013), *5-item Likert scale*	$\alpha = 0.85$
Information activities	Look up a definition of a word; get information about local events, find or check a fact; find information about about other people; investigate topics of personal interest; look for jobs, work; look for news; get information for school/college/homework; look for sports information; make travel plans (Q.C21a-e,i/Q.C29b,c,e in OxIS, 2013), *5-item Likert scale*	$\alpha = 0.84$
Entertainment activities	Play online games; get jokes, cartoons or other humorous content, listen to music online; watch movies, films or videos online; watch TV programs on the Internet (Q.C11a,c,e,g/Q.C21f in OxIS, 2013), *5-item Likert scale*	$\alpha = 0.79$
Transaction activities	Use your bank's online services; pay bills; buy a product online; compare products and prices; check your email; find location of a house, office, store, restaurant (Q.C30b,d–f,j/Q.C8a in OxIS, 2013), *5-item Likert scale*	$\alpha = 0.80$
Internet centrality		
Frequency of use	All items included in "Online activities" (see above), measuring the frequency of those activities that the individual actually does (0 = never set to missing value) (Q.C8a-c,e-i/Q.C9a,c,d/Q.C11a,c,e,g/Q.C21a-f,i/Q.C29b,c,e/Q.C30b,d,e,f,j in OxIS, 2013), *5-item Likert scale*	
Breadth of use	All items included in "Online activities" (see above), counting the those activities that the individual actually does, but not their frequency (all values set to 0 and 1) (Q.C8a-c,e-i/Q.C9a,c,d/Q.C11a,c,e,g/Q.C21a-f,i/Q.C29b,c,e/Q.C30b,d–f,j in OxIS, 2013), *continuous*	
Number of devices with Internet access	Does your household have: A hand-held tablet with a touch screen (iPad) / a hand held reader for books/mags (Kindle/Nook) / a TV with a built-in connection to the Internet; Do you yourself have a mobile phone?; How many working computers are available for people to use in this household? (Q.H7f,g,i/Q.H9/Q.H11 in OxIS, 2013), *continuous*	

Analytical concept	Quantitative representation	Validity
Internet centrality *(continued)*		
Number of places	Do you currently access the Internet: At home / travelling (mobile/wireless dongle) / at another person's home / at work / at school or university / at a public library? (Q.C1a–e,g in OxIS, 2013), *5-item Likert scale*	
Internet importance	How important is the Internet: For information, for entertainment (Q.A2c/Q.A3d in OxIS, 2013), *5-item Likert scale*	$\alpha = 0.89$
	"Internet information importance" (used in Chapter 6, subsection 6.1.1) consists of only the information variable (Q.A2c in OxIS, 2013), divided by the mean of the information importance of newspapers, radio and television (Q.A2a,b,d in OxIS, 2013)	
Internet experience	How many years have you been going online in total? (Q.C3n in OxIS, 2013), *free text answer*	
Reasons for non-use		
Skills	*Non-users* – too difficult to use; do not know how to use the Internet (Q.N1d,k in OxIS, 2013), *dichtomous*	$\alpha = 0.65$
	Ex-users – too difficult to use; never really figured out how to use (Q.E4d,k in OxIS, 2013), *dichtomous*	$\alpha = 0.81$
Devices	*Non-users* – no connection/computer available; it's too expensive (Q.N1b,c,f in OxIS, 2013), *dichtomous*	$\alpha = 0.67$
	Ex-users – no longer have a computer available (Q.E4c in OxIS, 2013), *dichtomous*	
Interest	*Non-users* – just not interested; it's not for people like me; it's not for people of my age (Q.N1a,m,n in OxIS, 2013), *dichtomous*	$\alpha = 0.60$
	Ex-users – just not interested; it's not for people like me; it's not for people of my age (Q.E4a,m,n in OxIS, 2013), *dichtomous*	$\alpha = 0.68$
Usage opportunities	*Non-users* – it's not useful; nothing of interest on the Internet (Q.N1e,j in OxIS, 2013), *dichtomous*	$\alpha = 0.68$
	Ex-users – it's not useful (Q.E4e in OxIS, 2013), *dichtomous*	
Most important non-use reason	And which of these reasons is the most important? Same grouping as above for "reasons of non-use" (Q.E5/Q.N2 in OxIS, 2013)	

Internet-based effects

Analytical concept	Quantitative representation	Validity
Internet outcomes		
Improved my health online	Have you ever found information that helped you improve your health? (Q.C22d in OxIS, 2013), *dichotomous*	
Saved money online	Have you ever saved money buying something online? (Q.C22b in OxIS, 2013), *dichotomous*	
Found a job	Have you ever found a job through the Internet? (Q.C22a in OxIS, 2013), *dichotomous*	
Learned a work-related skill	Have you ever used the Internet to learn new work-related skills? (Q.O13a in OxIS, 2013), *dichotomous*	
Social consequences		
Strengthens local community	Agreement: The Internet affected communities: Strengthened by social interaction" (Q.B3b in OxIS, 2013), *5-item Likert scale*	
Undermines local community	Agreement: The Internet affected communities – it undermined by isolating people (Q.B3a in OxIS, 2013), *5-item Likert scale*	
Allows me to keep in touch	Agreement: Going online allows me to keep in touch with people (Q.I8a in OxIS, 2013), *5-item Likert scale*	
Makes me feel less lonely	Agreement: When I'm online I don't feel lonely (Q.I8h in OxIS, 2013), *5-item Likert scale*	
Leads to addictive use	Agreement: I feel addicted to the Internet (Q.I2e in OxIS, 2013), *5-item Likert scale*	

Health

Analytical concept	Quantitative representation	Validity
Perceived health	How is your health in general? (Q.13 in ONS, 2011a), *5-item scale (Very good, good, fair, bad, very bad), recoded to 3-item scale (very good/good, fair, bad/very bad) for cross-tabulated data*, see Appendix VI (p. 306)	
Health service use	Total admissions per month (count of unique HES identification numbers) for hospital inpatient stays, *continuous variable divided by number of individuals in age/gender group in OA*, see Appendix VI (p. 307)	
Health-related use barriers	Does this health problem or disability limit your use of computers/the Internet? (Q.D17 in OxIS, 2013), *dichotomous*	

Antecedent factors

The 8 clustering dimensions used for the typology of health information seekers (Chapter 4) are shaded with light grey stripes (▨). The enablers "Information-finding ability" and "Information-evaluation ability" have not been included as clustering dimensions due to their similarity to general Internet skills, and as the latter is only asked for users (while the typology includes non-users too).

Analytical concept	Quantitative representation	Validity
Motivations		
Learning attitude	Agreement: I dislike learning new things*; I find it easy to understand new things; I learn new things on my own just because I want to; I am no good at learning new things*; I like any opportunity to discover something new (Q.I1e,g,i/SC.5h,i in OxIS, 2013), *5-item Likert scale*	$\alpha = 0.75$
Online enjoyment	*Users* – Agreement: Going online helps me pass time when I am bored; I just enjoy being online to see what comes up; When I'm online I don't feel lonely (Q.I8g,h,j in OxIS, 2013), *5-item Likert scale* *Non-users* – Agreement: The Internet is frustrating to work with* (Q.I2b in OxIS, 2013), *5-item Likert scale*	$\alpha = 0.76$
Trust in doctors	How much trust do you have in: Most doctors (Q.A7b in OxIS, 2013), *5-item Likert scale*	
Information usefulness	Agreement: The Internet makes life easier; Going online is an efficient means for finding information; The Internet helps me save time (Q.I2a,d/Q.I8c in OxIS, 2013), *5-item Likert scale*	$\alpha = 0.69$
Enablers		
Internet interest	Generally speaking, how interested would you say you are in the Internet? (Q.A9 in OxIS, 2013), *5-item Likert scale*, coded as 1 if "I am just not interested" was among reasons for non-use (Q.N1a in OxIS, 2013)	
Internet skills	How would you rate your ability to use the Internet? (Q.C6 in OxIS, 2013), *5-item scale*, missing values replaced based on "How confident are you to remove a virus?" and "I get nervous using technologies, because I might break something"* (Q.I1c/Q.C23c in OxIS, 2013); coded as 1 if "It's too difficult to use" or "Do not know how to use the Internet" were among reasons for non-use (Q.N1d,k in OxIS, 2013)	
Technology attitude	Agreement: Technology is making things better for people like me; I find it difficult to keep up to date with new technology*; It is a good idea to try new technologies; Often it is easier to do things without using technologies*; I do not trust technology, it fails when you need it the most*; I get nervous using technologies, because I might break something* (Q.B1a,c,g/Q.I1a-c in OxIS, 2013), *5-item Likert scale*	$\alpha = 0.81$
Self-efficacy	Agreement: What happens to me is my own doing (SC.5f in OxIS, 2013), *5-item Likert scale*	
Information-finding ability	Agreement: I find/found dealing with the amount of information online exhausting; I waste(d) time on irrelevant info to get what I looked for; I waste(d) too much time online; (Q.I2g,h/Q.I8f in OxIS, 2013), *5-item Likert scale*	$\alpha = 0.64$
Information-evaluation ability	How confident are you to judge the reliability of online content? (Q.C23a in OxIS, 2013) *4-item scale*	

Individual factors

The variables used for SM discussed in Chapter 3 are shaded in light grey (▢).

Analytical concept	Quantitative representation	Validity
Gender	[Interviewer observation] (Q.D2 in OxIS, 2013) **Matched with (see Appendix V, p. 303):** – Male – Female	
Age	In what year were you born? (Q.D1 in OxIS, 2013) *free text answer, transformed to age* **Matched with (see Appendix V, p. 302):** What is your date of birth? (Q.3 in ONS, 2011a), *free text answer, transformed to age*	
NS-SeC	Variable according to NS-SeC specifications (ONS, 2014c) based on labour force status, employment hierarchy/size of organisation and occupation code (Q.01, Q.O4b, Q.O4c and Q.O4d in OxIS, 2013) **Matched with (see Appendix V, p. 304):** Variable NS-SeC3 based on labour force status, employment hierarchy/size of organisation and occupation code (Q.30–Q.38 in ONS, 2011a)	
Education	What is your highest educational or vocational qualification? (Q.L1 in OxIS, 2013), *details on all answer choices included in Appendix V* **Matched with (see Appendix V, p. 303):** Which of these qualifications do you have? (Q.25 in ONS, 2011a), *details on all answer choices included in Appendix V*	
Long-term condition	Do you have a health problem or disability? (Q.D16 in OxIS, 2013) **Matched with (see Appendix V, p. 303):** Are your day-to-day activities limited because of a health problem or disability which has lasted, or is expected to last, at least 12 months? Include problems related to old age. (Q.23 in ONS, 2011a) – Yes, limited a lot – Yes, limited a little – No	
Lifestage	Which of the following describes your current situation? (Q.O1 in OxIS, 2013); Do any children (people under 18) live in your household? (Q.D5 in OxIS, 2013) – Student – Employed/unemployed **[AND]** no children – Employed/unemployed **[AND]** children – Retired	

Contextual factors

The variables used for SM discussed in Chapter 3 are shaded in light grey ( ).

Analytical concept	Quantitative representation	Validity
Social environment		
Fam./friends far away: call	Contact family/friend far away: Calling them on the phone? (Q.B5b in OxIS, 2013), *Likert scale*	
Fam./friends nearby: call	Contact family/friend nearby: Calling them on the phone (Q.B4b in OxIS, 2013), *Likert scale* Both call variables were set to 0 (no use) if "How often do you make or receive phone calls over the Internet?" was 0 (Q.C8e in OxIS, 2013)	
Fam./friends far away: email	Contact family/friend far away: Email or instant message them (Q.B5c in OxIS, 2013), *Likert scale*	
Fam./friends nearby: email	Contact family/friend nearby: Email or instant message them (Q.B4c in OxIS, 2013), *Likert scale*	
Fam./friends far away: visit	Contact family/friend far away: Going to visit them or they come here (Q.B5a in OxIS, 2013), *Likert scale*	
Fam./friends nearby: visit	Contact family/friend nearby: Going to visit them or they come here (Q.B4a in OxIS, 2013), *Likert scale*	
Met others online	Have you met someone on the Internet through: chat room, online community (hobby/interest group), discussion group or bulletin board, a message on a public board? (Q.C15c,d,i,j in OxIS, 2013), *count*	
Social embeddedness	How often do you feel you lack companionship; how often do you feel left out; how often do you feel isolated from others? (SC.3a-c in OxIS, 2013), *5-item Likert scale*	$\alpha = 0.91$
Help availability	How easily could you find someone to help you use the Internet? (Q.C5 in OxIS, 2013), *4-item scale*	
Medical expertise in household	Household Standard Occupation Code (2 digits): 22 – Health professionals **[AND]** 32 – Health/social care associate professionals (Q.O4d/Q.O6d in OxIS, 2013), *5-item Likert scale*	
Geographic environment		
Output area	OA code, 2001 census (OxIS, 2013) linked to OA in census (ONS, 2011a)	
Area classification	Area classifications cluster OAs based on the main socio-demographic, economic, and housing-related characteristics common to their population (Census, 2011k)	
Region	One of the 9 regions in England: London, South West, South East, East of England, West Midlands, East Midlands, North West, North East, Yorkshire and the Humber (Census, 2011l) *Note: Area classification/ region were summarised as "locality" for constraints (see Appendix V, p. 305)*	
Urbanity	Urban/rural characteristics of OA. Urban: conurbation, city and town. Rural: town and fringe, village, hamlets and isolated dwellings (Census, 2011j)	

V Constraint variables for spatial microsimulation

The following constraints were selected and entered into the model in reverse order: age, education, long-term health condition, gender, socio-economic status (NS-SeC), and locality. Ethnicity was excluded due to its low explanatory value. Following Anderson (2013), the constraints for SM must:

1. be correlated with individual variables taken to the small area level
2. be correlated with the target variables evaluated at the small area level
3. exist in both census and survey sharing the same definition/categories

To satisfy **conditions 1 and 2**, the table below shows the β coefficients resulting from a regression for all of the individual constraints and the cumulative Pseudo-R^2 following Cragg and Uhler (for details on locality see p. 305). For the aggregated target variables, the table shows Goodman-Kruskal's γ with significance based on the χ^2-test (although with large data, these are likely to be significant). Significance codes are $*** < 0.001; ** < 0.01; * < 0.05$.

	Individual variables				**Target variables**	
	Internet use		Health information seeking		Perceived health	Health service use
Independent variables	β	R^2_{cum}	β	R^2_{cum}	γ	γ
Age	-1.34 ***	0.34	0.09	0.09	0.577 ***	0.457 ***
Education	0.87 ***	0.48	0.47 ***	0.21	-0.451 ***	–
Long-term health condition	-0.25 ***	0.50	0.08	0.22	-0.941 ***	–
Gender	-0.14 *	0.50	0.29 ***	0.24	0.062 ***	0.152 ***
NS-SeC	-0.56 ***	0.52	-0.03	0.25	0.112 ***	–
Ethnicity – white	-0.07	0.52	0.12	0.25	–	–
– asian	-0.22		0.01		–	–
– black	0.03		0.03		–	–
N	**2,150**		**1,506**		**42,989,620**	**13,481,805**
Largest CI	**4.9**		**4.4**		–	–

To satisfy **condition 3**, the categories of each constraint are given below for OxIS and the census. HES is not included as it can be added to the census with a 1:1 merge based on OAs. Below are the categories with their coded values in the dataset, and their relative frequency as an indication of how well the definitions match.

OxIS		**Census**	
Age			
1 16–24	15.1%	**1** 16–24	14.6%
2 25–34	17.1%	**2** 25–34	16.7%
3 35–49	25.8%	**3** 35–49	26.3%
4 50–64	22.3%	**4** 50–64	22.3%
5 65–74	10.3%	**5** 65–74	10.6%
6 75–84	7.6%	**6** 75–84	6.8%
7 $\geq$ 85	1.8%	**7** $\geq$ 85	2.7%

OxIS		Census	
Education			
0 No qualifications	23.5%	**10** No academic or professional qualifications	22.5%
2 <=4 GCSE grade A–C **3** GCSE grade D–G **5** SNQ 4- grade 1–3 **7** <=4 O grade 1–3 **8** SNQ grades 4–7 **15** CSEs **17** <=4 O levels **22** GNVQ Foundation	17.8%	**11** Level 1: 1–4 O Levels/CSE/GCSEs (any grades), Entry Level, Foundation Diploma, NVQ Level 1, Foundation GNVQ, Basic/Essential Skills **16** Other: Vocational/work-related Qualifications, Foreign Qualifications (not stated/level unknown)	19.0%
1 >5 GCSE grade A–C **4** >5 SNQ grades 1–3 **6** >5 O grades 1–3 **9** SNQ Scottish Highers **11** SVQ Level 1/2 **16** >5 O levels **19** NVQ 1 or 2 **23** GNVQ Intermediate	18.6%	**12** Level 2: 5+ O Level (Passes)/CSEs (Grade 1)/GCSEs (Grades A*-C), School Certificate, 1 A Level/ 2–3 AS Levels/ VCEs, Intermediate/Higher Diploma, NVQ level 2, Intermediate GNVQ, City and Guilds Craft, BTEC First/General Diploma, RSA Diploma)	18.8%
10 SNQ Certificate Sixth Year Studies **12** SVQ level 3 **18** GCE A levels **20** NVQ 3 or 4 **24** GNVQ Advanced	13.1%	**13** Apprenticeship **14** Level 3: 2+ A Levels/VCEs, 4+ AS Levels, Higher School Certificate, Progression/Advanced Diploma, NVQ Level 3; Advanced GNVQ, City and Guilds Advanced Craft, ONC, OND, BTEC National, RSA Advanced Diploma	12.3%
14 SVQ 5 **21** NVQ 5 **25** Certificate/Diploma of Higher Education **26** HND **27** Bachelor's degree **28** Graduate Certificate/Diploma **29** Post-degree professional qualification **30** Master's degree **31** Doctoral degree	27.0%	**15** Level 4+: Degree (for example BA, BSc), Higher Degree (for example MA, PhD, PGCE), NVQ Level 4-5, HNC, HND, RSA Higher Diploma, BTEC Higher level, Foundation degree (NI), Professional qualifications (for example teaching, nursing, accountancy) **Source of summarised categories:** ONS (2014a), including details on all qualifications (not included in list of abbreviations)	27.4%
Long-term health condition			
0 No	16.7%	**1** Day-to-day activities not limited	79.1%
1 Yes	83.3%	**2** Day-to-day activities limited a little **3** Day-to-day activities limited a lot	20.9%
Gender			
0 Male	49.9%	**0** Male	49.2%
1 Female	50.1%	**1** Female	50.8%

OxIS		Census	
Socio-economic status (NS-SeC)			
1 NS-SeC3: Level 1	27.1%	**1/2** NS-SeC3: Level 1 – Higher managerial, administrative and professional occupations	30.7%
2 NS-SeC3: Level 2	22.5%	**3/4** NS-SeC3: Level 2 – Intermediate occupations	22.4%
3 NS-SeC3: Level 3	35.1%	**5/6/7** NS-SeC3: Level 3 – Routine and manual occupations	32.9%
.u NS-SeC3: Long-term unemployed	5.7%	**8** NS-SeC3: Never worked and long-term unemployed	5.9%
.o Others	9.6%	**9** Others: Students, occupations not stated or inadequately described, not classifiable for other reasons	8.2%

Below is the table used to assign **OxIS (2013)** respondents an NS-SeC level based on ONS (2014c). These refer to individual codes except for temporary homeworkers (e.g. looking after children), where the household codes were used.

OxIS: NS-SeC3 from Standard Occupational Classification (SOC2010) and employment status/organisation size (ONS, 2014c)	**Employment status/organisation size**						
	1 — Employers — large org.	**2** small org.	**3** Self-empl.	**4** — Managers — large org.	**5** small org.	**6** Super-visors	**7** Other employees
Occupation code							
11 Corporate managers and directors	1	2	2	1	1	1	1
12 Managers and proprietors	1	2	2	1	1	1	1
21 Science, research, engineering and technology professionals	1	1	1	1	1	1	1
22 Health professionals	1	1	1	1	1	1	1
23 Teaching and education professionals	1	1	1	1	1	1	1
24 Business, media, public service professionals	1	1	1	1	1	1	1
31 Science, engineering and technology associate professionals	1	1	1	1	1	1	1
32 Health/social care associate professionals	1	1	1	1	1	1	1
33 Protective service occupations	1	1	2	2	2	1	2
34 Culture, media and sports occupations	1	1	1	1	1	1	1
35 Business/public service associate professionals	1	1	1	1	1	1	1
41 Administrative occupations	1	2	2	2	2	2	2
42 Secretarial and related occupations	1	2	2	2	2	1	2
51 Skilled agricultural and related trades	1	2	2	3	3	3	3
52 Skilled metal, electrical and electronic trades	1	2	2	3	3	3	3
53 Skilled construction and building trades	1	2	2	3	3	3	3
54 Textiles, printing and other skilled trades	1	2	2	3	3	3	3
61 Caring personal service occupations	1	2	2	3	3	3	3
62 Leisure, travel and related personal service occupations	1	2	2	3	3	3	3
71 Sales occupations	1	2	2	2	2	2	2
72 Customer service occupations	1	2	2	2	2	2	2
81 Process, plant and machine operatives	1	2	2	3	3	3	3
82 Transport and mobile machine operatives	1	2	2	3	3	3	3
91 Elementary trades and related occupations	1	2	2	3	3	3	3
92 Elementary admin./ service occupations	1	2	2	3	3	3	3

Locality as a constraint is formed based on the ONS area classification (Census, 2011k), which comprises socio-economic, ethnic, work-related and urban/rural features of the area, as well as the region, which has shown to have differences in terms of Internet use (Blank et al., 2014) and health (British Academy, 2014; NHS England, 2010). Through including location as a constraint, the SM dataset of a certain OA was in fact replicated through individuals from a similar OA in terms of its area characteristics and its region, similarly to Anderson (2013).
However, to avoid the area constraint restricting diversity due to over-segmentation, similar regions (South, Midlands, North) and similar areas (for example, cosmopolitans, ethnicity central and multicultural metropolitans) were summarised as follows:

Locality (in the narrow sense)

	Regions								
Area classification	London	South East	South West	East of E.	East Midl.	West Midl.	York-shire	North East	North West
1 Rural residents	266				88		111		
2 Cosmopolitans **3** Ethnicity central **4** Multicult. metropolitans	243	128							
5 Urbanites	255				54		64		
6 Suburbanites	212				117		234		
7 Constrained city dwellers	97				108		203		
8 Hard-pressed living									

Where the OA could not be replicated based on these narrow constraints, the region constraint was released, so that only the area classification was used:

Locality (in the wide sense)

	Regions								
Area classification	London	South East	South West	East of E.	East Midl.	West Midl.	York-shire	North East	North West
1 Rural residents	465								
2 Cosmopolitans **3** Ethnicity central **4** Multicult. metropolitans	371								
5 Urbanites	343								
6 Suburbanites	563								
7 Constrained city dwellers	408								
8 Hard-pressed living									

VI Construction of outcome variables

These example calculations relate to the same example OA in Oxfordshire used in subsection 3.2.2. There are 300 individuals in the OA, but only 246 thereof are above the age of 16 (and thereby included in this research).

Perceived health

Example calculation for a 40-year-old female with education level 4 and NS-SeC 2, no long-term conditions.

LC3601EW: Perceived health by NS-SeC (Census, 2011d)

NS-SeC = *(all)*

H_{all}	H_1	H_2	H_3	H_4	H_5
246	4	14	49	77	102

$p_{1/2} = 0.29$ $p_{4/5} = 0.76$

Proportions used to split up education and age/gender/long-term health condition (released recoded 1–3 only)

NS-SeC = *2*

H_{all}	H_1	H_2	H_3	H_4	H_5
52	1	6	15	14	16

$= (1 \cdot H_1 +$
$2 \cdot H_2 +$
$3 \cdot H_3 +$
$4 \cdot H_4 +$
$5 \cdot H_5)/H_{all}$

$= (1.00 +$
$12.00 +$
$45.00 +$
$56.00 +$
$80.00)/52$

$=$ **3.73**

CT0465: Perceived health by education (Census, 2011c)

Education = *4*

H_{all}	$H_{1/2}$	H_3	$H_{4/5}$
64	1	5	58

$= (1 \cdot H_{1/2} \cdot p_{1/2} +$
$2 \cdot H_{1/2} \cdot (1 - p_{1/2}) +$
$3 \cdot H_3 +$
$4 \cdot H_{4/5} \cdot p_{4/5} +$
$5 \cdot H_{4/5} \cdot (1 - p_{4/5}))/H_{all}$

$= (0.29 +$
$1.43 +$
$15.00 +$
$175.16 +$
$71.05)/64$

$=$ **4.11**

LC3302EW: Perceived health by age, gender, long-term cond. (Census, 2011b)

Age = *34–49*
Gender = *female*
Long-term condition = *no*

H_{all}	$H_{1/2}$	H_3	$H_{4/5}$
29	0	1	28

$= (1 \cdot H_{1/2} \cdot p_{1/2} +$
$2 \cdot H_{1/2} \cdot (1 - p_{1/2}) +$
$3 \cdot H_3 +$
$4 \cdot H_{4/5} \cdot p_{4/5} +$
$5 \cdot H_{4/5} \cdot (1 - p_{4/5}))/H_{all}$

$= (0.00 +$
$0.00 +$
$3.00 +$
$84.56 +$
$34.30)/29$

$=$ **4.20**

Final: 4.01

Health service use

Example calculation for a 40-year-old female with education level 4 and NS-SeC 2, no long-term conditions.

HES data: Extract NIC-317612-V1Q3X (HES, 2013), count of inpatient HES-IDs		**LC1117EW**: Gender by age (Census, 2011i)
Gender = *female*		Gender = *female*
Age = *34–49*		Age = *34–49*
April 2013	0	No. of individuals in OA (male, 34–49): **30**
May 2013	0	
June 2013	0	
July 2013	0	
August 2013	0	
September 2013	1	
October 2013	2	
November 2013	0	
December 2013	1	
January 2014	0	
February 2014	0	
March 2014	1	
	5	**5/30 =** **Final: 0.16**

The resulting number ranges from 0 (nobody was in hospital) to 12 (everybody was in hospital at least once per month).

If the number of available individuals in the age/gender group in the OA was 0, the overall value was set to 0. This can occur in cases where individuals moved to a certain OA between 2011 and 2013 and then went to hospital.

For the same reason, in a few cases the final number is larger than 12: if individuals moved to an area between 2011 and 2013, they are included in HES data already, but not yet in the census. In these cases, the final number was capped at 12.

General inpatient data (main dataset): No ICD restrictions (all data was used).

Inpatient data for mental health conditions: ICD codes F01–F99 only.

Inpatient data for stigmatised health conditions: ICD codes F01–F99 (mental health), B20–24, R75, Z21 (HIV), A50–A60, A63, A64 (sexually transmitted diseases), N39 (incontinence and other disorders of the urinary system) only.

VII Internal validation results

Both tables include the mean values after 10 iterations [minimum; maximum].

	Pearson's R	**χ^2 test**[13]	**TAE**	**SAE**
Constraints	*corr. coefficient*	*p-value (χ^2,df)*	*absolute*	*relative (%)*
Age	0.99125 [0.99123; 0.99128]	1.00 (12,914; 155,100)	1,867,079 [1,863,624; 1,869,438]	4.34 [4.34; 4.35]
Education	0.97466 [0.97464; 0.97467]	1.00 (23,930; 171,050)	1,886,897 [1,884,528; 1,888,748]	4.39 [4.38; 4.39]
Long-term condition	0.99482 [0.99480; 0.99485]	1.00 (7,380; 147,741)	578,744 [576,856; 580,280]	1.60 [1.59; 1.60]
Gender	0.94934 [0.94931; 0.94940]	1.00 (35,552; 171,371)	1,078,102 [1,075,790; 1,080,118]	2.51 [2.50; 2.51]
NS-SeC	0.97675 [0.97674; 0.97675]	1.00 [29,207; 162,166]	1,809,842 [1,808,106; 1,811,874]	4.21 [4.21; 4.21]
Area (wide sense)	0.99999 [0.99999; 0.99999]	1.00 (34; 34,273)	708 [700; 718]	0.00 [0.00; 0.00]
Area[14] (narrow sense)	0.99234 [0.99234, 0.99234]	0.93 (1,085; 12,240)	1,546,136 [1,546,136; 1,546,136]	3.60 [3.60;3.60]
Overall value	**0.98400** [0.98400, 0.98401]	**1.00** (17,002; 136,982)	**7,221,372** [7,214,932; 7,227,372]	**17.05** [17.03; 17.06]

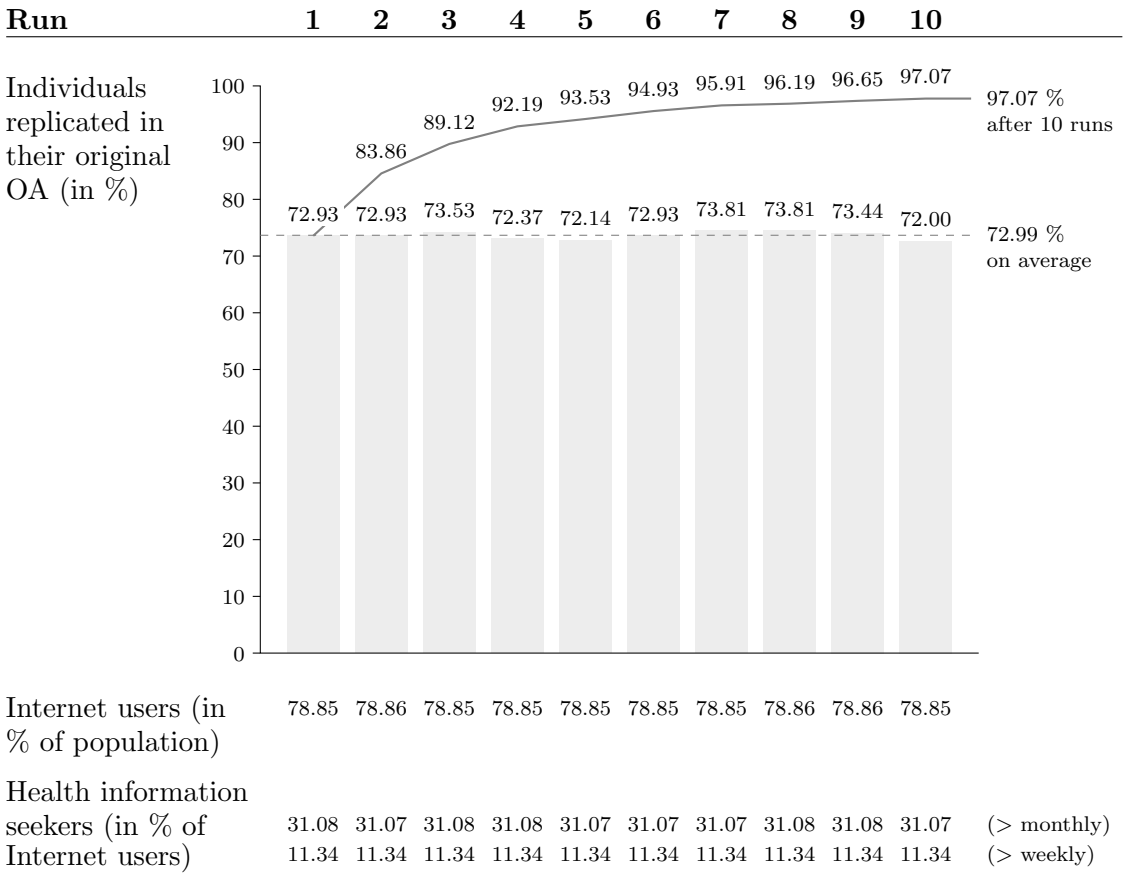
	Deviation > 5%	**Deviation > 5**	**RMSE**	**NRMSE**
Constraints	*relative (%)*	*relative (%)*	*absolute*	*relative*
Age	41.44 [41.35; 41.50]	2.31 [2.30; 2.33]	2.17 [2.16; 2.17]	0.06 [0.06; 0.06]
Education	36.76 [36.67; 36.77]	5.67 [5.65; 5.71]	5.25 [5.25; 5.26]	0.20 [0.20; 0.20]
Long-term condition	17.10 [16.97; 17.12]	3.87 [3.80; 3.93]	2.63 [2.62; 2.63]	0.23 [0.23; 0.23]
Gender	8.31 [8.24; 8.37]	9.70 [9.58; 9.81]	13.50 [13.49; 13.50]	0.64 [0.64; 0.64]
NS-SeC	40.00 [39.94; 40.06]	5.12 [5.09; 5.14]	6.74 [6.74; 6.74]	0.17 [0.17; 0.17]
Area (wide sense)	0.00 [0.00; 0.00]	0.00 [0.00; 0.00]	0.44 [0.44; 0.45]	0.01 [0.01; 0.01]
Area[14] (narrow sense)	0.51 [0.51; 0.51]	0.51 [0.51; 0.51]	10.51 [10.51; 10.51]	0.26 [0.26; 0.26]
Overall value	**27.87** [27.83; 27.88]	**3.74** [3.72; 3.76]	**30.73** [30.72; 30.74]	**1.31** [1.31; 1.31]

[13] Degrees of freedom differ per constraint as cell frequencies < 5 were eliminated.

[14] Not included in overall goodness-of-fit statistic (counting area constraint only once).

Replication of individuals in their original OAs

The chart below shows the proportion of individuals in the simulated dataset which was replicated in the OA where the original OxIS respondent came from. This proportion is given for 10 datasets, each of which has been obtained with a different seed value, thereby altering the basis of the probabilistic element of the TRS algorithm (Lovelace and Ballas, 2013).
The chart shows that, on average, 73.0% of individuals in the simulated dataset are replicated at least once in their original OA (obtained by filtering unique serial numbers matching the serial numbers in OxIS for each of the 260 OAs). Aggregating several simulated datasets shows that no individuals were systematically excluded, but slightly different individuals were chosen each time, so that the total coverage asymptotically approached 100%. Aggregating 10 simulations, 97.1% of individuals were replicated in their OA.



However, this did not greatly alter the key variables of interest coming from these individuals (for example, Internet use and health information-seeking). As a consequence, the dataset of the first run (seed = 68,649) has been used throughout the analyses.

VIII External validation results

Overview of external validation datasets

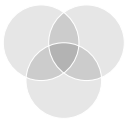
Black-framed concepts are the main outcome variables. Concepts in light grey are only partially covered, while some aspects of importance to the analysis are missing.



	SM dataset	Opinions & Lifestyle Survey (OLS)	Adult Media Use Survey (AMU)	UK Household Longit. survey (UKHLS)
Source	OII, ONS, HSCIC	Office of National Statistics (ONS)	Office of Communications (OfCom)	University of Essex Institute for Social Economic Research
Method	Simulated dataset	Survey	Survey	Survey
Collection	Offline	Offline	Offline	Offline
Sampling	OxIS: Multi-stage random sample, ≥16 years in households	Multi-stage stratified random sample, ≥16 years in households	Random sample, ≥16 years in households	Multi-stage stratified random sample, ≥16 years in households
Geography	England	Great Britain	UK	UK
Fieldwork times	February 2013 to April 2013	January 2013 to March 2013	October 2012 to November 2012	January 2012 to May 2014
Individuals England, ≥ 16	2,150	2,520	1,321	36,254
Internet use	User concepts	User concepts	User concepts	User concepts
	Online activities	Online activities	Online activities	Online activities
	Internet centrality	Internet centrality	Internet centrality	
	Reasons for non-use	Reasons for non-use	Reasons for non-use	
Internet-based effects	Internet outcomes		Internet outcomes	
	Social consequences		Social consequences	
Health	Health use barriers	Health use barriers	Health use barriers	Health use barriers
	Perceived health	Perceived health		Perceived health
	Health service use			Health service use
Antecedent factors	Motivations		Motivations	
	Enablers	Enablers	Enablers	
Individual factors	Age	Age	Age	Age
	Gender	Gender	Gender	Gender
	NS-SeC	NS-SeC	NS-SeC	NS-SeC
	Education	Education	Education	Education
	Long-term condition	Long-term condition	Long-term condition	Long-term condition
	Life stage	Life stage	Life stage	Life stage
Contextual factors	Social environment		Social environment	Social environment
	Geographic env.	Geographic env.	Geographic env.	Geographic env.

Question wording for external datasets

Below are all the items used for external validation, as well as how they were recoded in order to make them as comparable as possible.



	SM dataset	Opinions & Lifestyle Survey (OLS)	Adult Media Use Survey (AMU)	UK Household Longit. Survey (UKHLS)
Internet users	Do you personally use the Internet at home/work/ school/college/ elsewhere? – Yes – No, but used it in the past – No, never *user/ non-user recoded*	When did you last use the Internet? – within the last 3 months – 3 months to 1 year ago – More than 1 year ago – Never used it	Item: "Go online at home with a PC/laptop/net-book or alter-native device", *created by OfCom through variables IN1–IN3* – Yes – No	How often do you use the Internet for your personal use? – Every day – Several times per week – Several times per month – Once a month – Less than once per month – Never use – No access at home, work or elsewhere
Health information seeking	How often do you find information about health or medical care? – > than daily – daily – weekly – monthly – < than monthly – never *weekly recoded* *monthly recoded*	In the last 3 months, for which of the following did you use the Internet? – Seeking health-related infor-mation	Activities online: Find information about health-related issues – > once a week – > every 3 weeks – less often – never do this	
Perceived health	How is your health in general? – Very good – Good – Fair – Bad – Very bad *recoded to 1-3 scale in census data*	How is your health in general? – Very good – Good – Fair – Bad – Very bad		In general, would you say your health is... – Excellent – Very good – Good – Fair – Poor
Health service use	Visits to hospital (in-/outpatient) within a year			Which of these services have you used over the last 12 months? – A local hospital

Comparison of simulated values with other survey data

External validation shows that the simulated dataset is an acceptable representation of the real world compared to other datasets. For the proportion of Internet users and health information-seekers, as well as for the average rating of perceived health, the simulated data falls into the **95% confidence intervals** of the external validation datasets in many cases. In all cases, the confidence intervals may be asymmetric as they are calculated using a logit transformation and Yates' (1934) continuity correction to account for the bounded nature of proportions (between 0 and 100%) and the perceived health rating (between 1 and 3 to harmonise the scales between the different datasets).

At the same time, there are some deviations between the simulated data and the external validation datasets, which may partially be explained by structural differences in the way the data was obtained:

- **Internet users** *(p. 313)*: While OxIS asked whether individuals "personally use the Internet on whatever device at home, work, school, college or elsewhere", AMU listed all devices (personal computer, laptop, netbook or alternative device), but restricted the question to the home. In contrast, the OLS and the UKHLS asked for general use without specifying devices or locations.
- **Health information-seekers** *(p. 313)*: OxIS and AMU asked about the frequency of looking for health information online (with "at least monthly" as one of the answers), whereas OLS asked about having looked for health-related content online in the last 3 months.
- **Perceived health** *(p. 314)*: The different scale for the UKHLS with 2 as the neutral option (5: Excellent; 4: Very good; 3: Good; 2: Fair; 1: Poor) may have elicited different responses. In addition, the surveys featured different proportions of respondents with long-term conditions: the census shows that about 21% of the English population has a long-term condition, but 35% of respondents in the OLS and the UKHLS do so. In OxIS, only 16% of respondents have a long-term condition.
- **Health service use** *(p. 315)*: As opposed to the simulated dataset, the UKHLS asked about health service use in the past 12 months (without frequencies), and did not specify whether inpatient and outpatient services should both be included in this answer. Also, the UKHLS only states whether health service use was mentioned in response to the question "Which of these services have you used over the last 12 months?"

Internet users (% of population)

Regions	*OxIS*	*Simulation*	*OLS*	*AMU*	*UKHLS*
London	85.5	83.3	91.4	78.2	86.5
confidence interval	[81.1; 89.0]	[83.3; 83.3]	[87.3; 94.3]	[71.6; 83.6]	[85.6; 87.4]
South East	82.9	80.0	89.0	79.2	87.1
confidence interval	[78.5; 86.6]	[80.0; 80.0]	[85.4; 91.8]	[72.6; 84.6]	[86.2; 88.0]
South West	81.2	77.9	86.9	79.3	82.5
confidence interval	[75.1; 86.1]	[77.9; 77.9]	[82.3; 90.5]	[71.3; 85.6]	[81.2; 83.7]
East of England	78.3	78.1	85.3	79.3	84.8
confidence interval	[72.5; 83.2]	[78.1; 78.1]	[80.5; 89.1]	[72.1; 85.1]	[83.6; 85.9]
East Midlands	83.2	79.6	85.8	76.7	82.2
confidence interval	[76.7; 88.1]	[79.6; 79.6]	[80.3; 90.0]	[67.9; 83.7]	[80.9; 83.4]
West Midlands	78.7	79.3	80.1	76.4	80.5
confidence interval	[72.7; 83.7]	[79.3; 79.3]	[74.4; 84.8]	[67.9; 83.3]	[79.2; 81.8]
Yorkshire	75.3	76.4	81.3	76.1	80.9
confidence interval	[68.9; 80.8]	[76.4; 76.4]	[76.1; 85.6]	[67.8; 82.9]	[79.6; 82.1]
North East	60.0	74.4	74.7	74.3	78.2
confidence interval	[50.0; 69.2]	[74.3; 74.5]	[66.1; 81.8]	[62.5; 83.5]	[76.2; 80.1]
North West	73.7	76.3	81.9	78.1	81.2
confidence interval	[68.1; 78.7]	[76.3; 76.3]	[77.6; 85.5]	[71.2; 83.8]	[80.0; 82.3]
Total	**79.1**	**78.8**	**85.1**	**77.5**	**83.3**
confidence interval	[77.3; 80.8]	[78.8; 78.8]	[83.6; 86.4]	[75.1; 79.7]	[82.9; 83.7]

Health information-seekers, monthly (% of Internet users)

Regions	*OxIS*	*Simulation*	*OLS*	*AMU*	*UKHLS*
London	43.9	37.3	54.2	–	–
confidence interval	[37.7; 50.3]	[37.3; 37.4]	[47.9; 60.4]	–	–
South East	26.0	26.8	46.2	–	–
confidence interval	[20.9; 31.9]	[26.8; 26.8]	[41.0; 51.5]	–	–
South West	33.7	28.2	46.3	–	–
confidence interval	[26.4; 41.9]	[28.2; 28.2]	[40.1; 52.7]	–	–
East of England	31.7	33.7	46.5	–	–
confidence interval	[24.9; 39.4]	[33.7; 33.7]	[40.1; 53.0]	–	–
East Midlands	24.5	27.7	46.4	–	–
confidence interval	[17.8; 32.7]	[27.6; 27.8]	[39.2; 53.8]	–	–
West Midlands	37.8	33.2	39.9	–	–
confidence interval	[30.3; 45.9]	[33.2; 33.2]	[33.0; 47.2]	–	–
Yorkshire	45.1	29.9	40.2	–	–
confidence interval	[36.9; 53.6]	[29.9; 30.0]	[33.8; 46.9]	–	–
North East	30.4	31.7	30.6	–	–
confidence interval	[19.4; 44.0]	[31.6; 31.8]	[21.8; 40.9]	–	–
North West	20.0	31.1	35.3	–	–
confidence interval	[14.6; 26.7]	[31.1; 31.1]	[30.1; 40.9]	–	–
Total	**32.7**	**31.1**	**44.0**	**31.0**	**–**
confidence interval	[30.4; 35.1]	[31.1; 31.1]	[41.9; 46.1]	[28.2; 33.9]	–

Perceived health: Comparison of average for users and non-users

Average rating on a scale from 1 to 3 (recoded to harmonise scales)

	Users			**Non-users**		
Group	*Simulation*	*OLS*	*UKHLS*	*Simulation*	*OLS*	*UKHLS*
Male:						
18-24	2.86	2.94	2.91	2.79	3.00	2.93
confidence interval	[2.86; 2.86]	[2.79; 2.99]	[2.89; 2.93]	[2.79; 2.79]	[1.40; 3.00]	[2.56; 3.00]
25-34	2.82	2.89	2.89	2.75	2.02	2.55
confidence interval	[2.82; 2.82]	[2.78; 2.95]	[2.87; 2.91]	[2.75; 2.75]	[1.30; 2.72]	[2.24; 2.76]
35-49	2.80	2.79	2.83	2.66	2.42	2.50
confidence interval	[2.80; 2.80]	[2.70; 2.86]	[2.81; 2.85]	[2.66; 2.66]	[2.00; 2.72]	[2.36; 2.62]
50-64	2.71	2.72	2.71	2.50	2.28	2.39
confidence interval	[2.71; 2.71]	[2.61; 2.80]	[2.68; 2.73]	[2.50; 2.50]	[1.95; 2.56]	[2.31; 2.47]
65-74	2.62	2.70	2.70	2.49	2.32	2.34
confidence interval	[2.62; 2.62]	[2.55; 2.81]	[2.66; 2.74]	[2.49; 2.49]	[2.05; 2.55]	[2.26; 2.41]
75-84	2.63	2.55	2.56	2.43	2.24	2.30
confidence interval	[2.63; 2.63]	[2.23; 2.77]	[2.48; 2.63]	[2.43; 2.43]	[1.96; 2.49]	[2.22; 2.38]
85+	2.48	2.69	2.49	2.38	2.28	2.28
confidence interval	[2.48; 2.48]	[1.93; 2.97]	[2.21; 2.70]	[2.38; 2.38]	[1.84; 2.63]	[2.13; 2.42]
Female:						
18-24	2.80	2.91	2.90	2.80	–	2.65
confidence interval	[2.80; 2.80]	[2.78; 2.97]	[2.88; 2.92]	[2.80; 2.80]	[–; –]	[2.28; 2.86]
25-34	2.86	2.89	2.85	2.73	3.00	2.58
confidence interval	[2.86; 2.86]	[2.80; 2.94]	[2.83; 2.87]	[2.73; 2.73]	[1.93; 3.00]	[2.36; 2.74]
35-49	2.83	2.81	2.81	2.68	2.61	2.38
confidence interval	[2.83; 2.83]	[2.74; 2.87]	[2.79; 2.83]	[2.68; 2.68]	[2.01; 2.90]	[2.28; 2.47]
50-64	2.80	2.77	2.73	2.66	2.63	2.32
confidence interval	[2.80; 2.80]	[2.68; 2.84]	[2.71; 2.75]	[2.66; 2.66]	[2.36; 2.81]	[2.25; 2.39]
65-74	2.69	2.74	2.73	2.59	2.44	2.47
confidence interval	[2.69; 2.69]	[2.59; 2.84]	[2.69; 2.77]	[2.59; 2.59]	[2.22; 2.62]	[2.41; 2.53]
75-84	2.69	2.59	2.65	2.45	2.25	2.42
confidence interval	[2.69; 2.69]	[2.30; 2.78]	[2.56; 2.73]	[2.45; 2.45]	[2.05; 2.44]	[2.35; 2.48]
85+	2.67	2.61	2.58	2.50	2.12	2.36
confidence interval	[2.67; 2.67]	[1.75; 2.96]	[2.26; 2.79]	[2.50; 2.50]	[1.80; 2.42]	[2.24; 2.47]
Total	**2.79**	**2.81**	**2.80**	**2.55**	**2.36**	**2.39**
confidence interval	[2.79; 2.79]	[2.78; 2.83]	[2.79; 2.81]	[2.55; 2.55]	[2.28; 2.44]	[2.36; 2.41]

Health service use: Comparison of average for users and non-users

% of individuals who used health services over the last 12 months

	Users		**Non-users**	
Group	*Simulation* inpatient[15]	*UKHLS* not specified	*Simulation* inpatient[15]	*UKHLS* not specified
Male:				
18-24	0.7	4.1	0.7	0.0
confidence interval	[0.7; 0.7]	[3.4; 5.2]	[0.6; 0.8]	[0.0; 0.9]
25-34	0.9	4.6	0.9	3.1
confidence interval	[0.9; 0.9]	[3.7; 5.6]	[0.8; 1.0]	[1.8; 5.1]
35-49	1.2	9.8	1.2	8.7
confidence interval	[1.2; 1.2]	[8.5; 11.2]	[1.2; 1.2]	[6.4; 11.6]
50-64	2.0	4.2	2.3	7.2
confidence interval	[2.0; 2.0]	[3.4; 5.2]	[2.2; 2.3]	[5.2; 9.9]
65-74	3.9	1.6	4.1	6.6
confidence interval	[3.9; 3.9]	[1.1; 2.2]	[4.1; 4.1]	[4.7; 9.3]
75-84	5.9	0.5	6.4	3.8
confidence interval	[5.8; 6.0]	[0.2; 0.9]	[6.4; 6.5]	[2.4; 6.0]
85+	10.8	0.0	10.9	1.2
confidence interval	[10.7; 10.9]	[0.0; 0.2]	[10.8; 11.0]	[0.5; 2.7]
Female:				
18-24	1.5	4.7	1.7	8.9
confidence interval	[1.5; 1.5]	[3.9; 5.7]	[1.6; 1.8]	[6.2; 12.5]
25-34	2.2	9.7	2.5	16.3
confidence interval	[2.2; 2.2]	[8.5; 11.0]	[2.5; 2.5]	[12.7; 20.7]
35-49	1.6	13.2	1.8	16.8
confidence interval	[1.6; 1.6]	[11.8; 14.7]	[1.8; 1.8]	[13.1; 21.2]
50-64	2.0	4.4	2.4	14.4
confidence interval	[2.0; 2.0]	[3.6; 5.4]	[2.4; 2.4]	[11.0; 18.6]
65-74	3.4	1.2	3.6	12.0
confidence interval	[3.4; 3.4]	[0.8; 1.8]	[3.6; 3.6]	[8.9; 15.9]
75-84	4.8	0.3	4.9	4.0
confidence interval	[4.7; 4.9]	[0.2; 0.7]	[4.9; 4.9]	[2.3; 6.8]
85+	6.5	0.1	6.7	1.1
confidence interval	[6.3; 6.7]	[0.0; 0.3]	[6.6; 6.8]	[0.4; 3.1]
Total[16]	**3.4**	**4.9**	**3.6**	**4.1**
confidence interval	[3.4; 3.4]	[4.6; 5.3]	[3.6; 3.6]	[3.5; 4.8]

[15] Divided by 12 to align the timeframe between the simulation and the UKHLS.

[16] The numbers do not correspond to Table 6, as they are split by users/non-users.

IX Spatial microsimulation in R: approach and code

All analyses were done using the statistical package R 3.1 with RStudio on Windows 8.1 (64-bit) and Linux Ubuntu 14.04. This appendix contains an overview of how the data was set up and manipulated in R as a way of making the source code accessible and to allow the results to be reproduced.

Overview of main stages and files

Stage	Filename	Short description
Preparation	Packages *(01_packages.R)*	Installs and loads all necessary packages and other global settings
	Survey *(02_prep_survey.R)*	Recodes all constraint variables (see p. 302) and further constructs (see p. 295) in OxIS
	Census *(03_prep_census.R)*	Recodes all constraint variables (see p. 302) and values for perceived health in the census
	HES data *(04_prep_hes.R)*	Recodes all values for health service use from HES data
Data formation	Constraint selection *(05_form_conssel.R)*	Runs regressions of potential constraints on individual and target variables of interest
	Microsimulation *(06_form_sm.R)*	Includes the spatial microsimulation (IPF procedure and TRS integerisation)
	Data addition *(07_form_dataadd.R)*	Adds all survey and census data to serial number/OA combination from previous step
	Combination *(08_form_comb.R)*	Combines all created datasets for evaluation and analysis
Validation	Internal validation *(09_valid_int.R)*	Includes all measures for internal validation (see p. 308)
	External validation *(10_valid_ext.R)*	Includes all measures for external validation from OLS, AMU and UKHLS (see p. 310)
Analysis	Analysis RQ_1 *(11_analysis_rq1.R)*	Includes all quantitative analyses shown in Chapter 4: Regressions, SEM, cluster analysis
	Analysis RQ_2 *(12_analysis_rq2.R)*	Includes all quantitative analyses shown in Chapter 5: SEM, multinomial regression
	Analysis RQ_3 *(13_analysis_rq3.R)*	Includes all quantitative analyses shown in Chapter 6: Regressions
Visualisation	Map creation *(14_vis_shapefiles.R)*	Includes sourcecode for maps produced based on shapefiles for Chapters 3 and 6

Details and code highlights

The code in this research is based on Lovelace (2014) obtained during a two-day workshop at the University of Leeds (8–9 May 2014). The mathematical notation builds on Lovelace (2014), Lovelace et al. (2015), Ballas and Anderson (2013) and Wong (1992).

Preparation

The most important packages used in the R code include the following:

- **ipfp** for iterative proportional fitting
- data manipulation (**Hmisc, plyr, data.table, doBy, foreign, matrixStats**)
- analysis and goodness-of-fit (**lavaan, nlme, nnet, psy, car, perturb, fpc**)
- efficient dealing with large datasets (**biglm, ff, ffbase**)
- visualisation and mapping (**shapefiles, maptools**)

The three files that recode survey, census and HES data do so using the standard commands in R. The construction of learning attitude is provided as an example of the survey constructs below:

From 02_prep_survey.R:

```
# Attitudes: Learning attitude
temp_learnatt.df <- as.data.frame(cbind(survey.df$aglerdis, # Build temp data frame
                                        survey.df$aglereasy,
                                        survey.df$aglerwant,
                                        survey.df$agnoler,
                                        survey.df$agopp))
temp_learnatt.df <- t(apply(temp_learnatt.df,1,function(x) x-1)) # make Likert scale
temp_learnatt.df[,c(1,4)] <- t(apply(temp_learnatt.df[,c(1,4)], # reverse items 1 and 4
                               1,function(x) x*(-1)+6))
cronbach(na.omit(temp_learnatt.df)) # test construct validity
int_use.df$att_learn <- rowMeans(temp_learnatt.df,na.rm=TRUE) #build final variable
sum(is.na(as.matrix(int_use.df$att_learn))) # test missing values
int_use.df$att_learn[is.na(int_use.df$att_learn)] <- 3 # recode missing data as neutral
int_use.df$att_learn[1:10] # check values
summary(int_use.df$att_learn) # check mean, median and range
```

The OAs in their 2001 definition in OxIS were mapped to OAs in 2011 definition based on the lookup table *OA01_OA11_LAD11_EW_LU* (Census, 2011m). For the one OxIS OA that was split between 2001 and 2011 into 4 areas with similar area characteristics, one of the four new OAs was selected.

An example of recoding of a constraint variable from the survey to the defined categories in Appendix V (p. 302) is provided below.

From 02_prep_survey.R:

```
# Build survey constraint 1: Age
survey.df$age7 <- cut(survey.df$age_new,
                      breaks=c(0,24,34,49,64,74,84,99),
                      labels=c("a24","a34","a49","a64","a74","a84","a99"))
cbind(survey.df[1:10,815:817],survey.df$age_new[1:10]) # inspect data
table(survey.df$age7) # check proportions
```

Using the analogous categories, the census was recoded based on the original comma-separated values files *(LC5102EW, LC3302EW, CT0465, LC3601EW, LC6121EW, QS103EW, LC5202EW)* as provided by the Census (2011a,b,c,d,f,g,h):

From 03_prep_census.R:

```
# Build census constraint 1: Age
census_int.df <- as.data.frame(census_raw3.df[,2])
names(census_int.df) <- c("oa11")
census_int.df$oa11 <- as.numeric(substr(census_int.df$oa11,2,9))
census_int.df[, "a24"] <- rowSums(census_raw3.df[,22:30])
census_int.df[, "a34"] <- rowSums(census_raw3.df[,31:40])
census_int.df[, "a49"] <- rowSums(census_raw3.df[,41:55])
census_int.df[, "a64"] <- rowSums(census_raw3.df[,56:70])
census_int.df[, "a74"] <- rowSums(census_raw3.df[,71:80])
census_int.df[, "a84"] <- rowSums(census_raw3.df[,81:90])
census_int.df[, "a99"] <- rowSums(census_raw3.df[,91:106])
census.df <- merge(census.df,census_int.df,all=FALSE)
census.df$total[1:10] == rowSums(census.df[1:10, 45:51]) # check if TRUE
```

In addition, the census preparation also consisted of creating the lookup matrix for perceived health and health service use based on the available cross-tabulations published by ONS. For the example of perceived health, a first matrix was created based on the cross-tabulation of age, gender and long-term conditions from the original csv file *LC3302EW* (Census, 2011b), as shown below. As perceived health was recoded to 1–3 due to statistical disclosure control, the proportions between good/very good and bad/very bad were obtained separately from *LC3601EW* (Census, 2011d) to reconstruct the 1–5 scale as closely as possible. A second matrix (not shown in code) was created analogously for education from *CT0465* (Census, 2011c), and a third one for NS-SeC from *LC3601EW* (Census, 2011d). For the final lookup matrix used in *07_form_dataadd.R*, all three matrices were multiplied.

From 03_prep_census.R:

```
# Obtain proportions between bad/very bad and good/very good
prop.df <- as.data.frame(cbind(census_raw5.df[,20]/rowSums(census_raw5.df[,c(20,36)]),
           census_raw5.df[,84]/rowSums(census_raw5.df[,c(68,84)])))
prop.df <- cbind(census_raw5.df[,2],prop.df)
names(prop.df) <- c("oa11","prop_good","prop_bad")
prop.df[is.na(prop.df)] <- 0.5 # if both are 0 replace with equal proportions

# Build perceived health lookup table (part 1/3)
health.df <- as.data.frame(census_raw4.df[,2])
names(health.df) <- c("oa11")

# Compute average score for perceived health for age, gender and disability
a <- 1
i <- 184
for(b in 1:14){
  # For males
  health.df[,paste('health',toString(colnames(cons)[27+a]),
  toString(colnames(cons)[31]),toString(colnames(cons)[41+b-7*(a-1)]),sep=".")] <-
     ((census_raw4.df[,i+1]+census_raw4.df[,i+5])*census_raw4.df[,"prop_good"]*5+
     (census_raw4.df[,i+1]+census_raw4.df[,i+5])*(1-census_raw4.df[,"prop_good"])*4+
     (census_raw4.df[,i+2]+census_raw4.df[,i+6])*3+
     (census_raw4.df[,i+3]+census_raw4.df[,i+7])*(1-census_raw4.df[,"prop_bad"])*2+
     (census_raw4.df[,i+3]+census_raw4.df[,i+7])*census_raw4.df[,"prop_bad"]*1)) /
     (census_raw4.df[,i]+census_raw4.df[,i+4])
```

```
  # For females
  health.df[,paste('health',toString(colnames(cons)[27+a]),
  toString(colnames(cons)[30]),toString(colnames(cons)[41+b-7*(a-1)]),sep=".")] <-
     ((census_raw4.df[,i+1]+census_raw4.df[,i+9])*census_raw4.df[,"prop_good"]*5+
     (census_raw4.df[,i+1]+census_raw4.df[,i+9])*(1-census_raw4.df[,"prop_good"])*4+
     (census_raw4.df[,i+2]+census_raw4.df[,i+10])*3+
     (census_raw4.df[,i+3]+census_raw4.df[,i+11])*(1-census_raw4.df[,"prop_bad"])*2+
     (census_raw4.df[,i+3]+census_raw4.df[,i+11])*census_raw4.df[,"prop_bad"]*1)) /
     (census_raw4.df[,i]+census_raw4.df[,i+8])
  i <- i+16
  if (b==7) {
     a <- 2
  }
}
```

In contrast, HES inpatient data was transformed by a simple aggregation, as it was already requested in a format required for this research (OA 2001, OA 2011, age group, gender, diagnosis (3-digit), month, count of events, count of HES-IDs):

From 04_prep_hes.R:

```
# HES-IDs: Admitted patient care
hsu.df <- as.data.frame(hsu.df[, sum(count_hesids,
    na.rm=TRUE),by=list(hsu.df$oa11,hsu.df$age7,hsu.df$gender)])
hsu.df <- as.data.frame.matrix(hsu.df)
colnames(hsu.df) <- c("oa11","age7","gender","count_hesids")
```

For analysing stigmatised and mental health conditions, the data was restricted to selected ICDs (see Appendix VI, p. 307) before running the code above.

Data formation

First, the constraints were selected in *05_form_conssel.R*. Selection was based on a (logistic) regression for Internet use and health information seeking as the two main survey concepts, and Goodman-Kruskal's γ and a chi-square test for perceived health and health service use as the outcome concepts (see Appendix V, p. 302).

The next step then consisted of determining the weights in *06_form_sm.R* as described in (Lovelace, 2014; Lovelace et al., 2015). The weights $w(t)_{ija}$ are determined with an iterative algorithm from t = 1, 2, 3, ..., n, which converges to a solution at iteration t with respect to how often a certain individual from the survey needs to be selected in order to replicate a specific output area $a = \{1, 2, 3, ..., 171.372\}$ based on the chosen constraints. For simplicity, there are only two constraints, i and j (for example, gender and long-term condition). From the census data, we obtain $\sum_j C_{ija}$ and $\sum_i C_{ija}$ as the total of individuals in one area for a certain category of the constraint variable, for example $\sum_j C_{i=1;ja}$ as the sum of females in an area a.

At $t = 1$, each individual has a starting weight of $w(1)_{ija} = 1$. The weights $w(t+1)_{ija}$ for an individual in one area a with certain characteristics of i and j (a specific category of gender and long-term condition, for example female and no long-term condition) at iteration $t+1$ can then be calculated as follows:

$$w(t+1)_{ija} = \begin{cases} w(t)_{ija} \cdot \frac{\sum_i C_{ija}}{\sum_i w(t)_{ija}} \text{ for } t = \{0, 2, 4, ..., n\} \\ w(t)_{ija} \cdot \frac{\sum_j C_{ija}}{\sum_j w(t)_{ija}} \text{ for } t = \{1, 3, 5, ..., n+1\} \end{cases} \quad (1)$$

In each iteration t, this algorithm adjusts the weights for the individual so that the marginal totals for constraint i are adjusted as closely as possible (often referred to as the row margins) in order to calculate $w(t+1)_{ija}$, which is then used as input for the next iteration where the marginal total for constraint j (the column totals) are adjusted in order to calculate $w(t+2)_{ija}$. In theory, these algorithms converge at $t = n$, when the sum of weights $w(n)_{ija}$ is equal to both the census constraints $\sum_j C_{ija}$ and $\sum_i C_{ija}$ for all categories of i and j respectively (Wong, 1992):

$$\sum_i w(n)_{ija} = \sum_i C_{ija} \quad \text{and} \quad \sum_j w(n)_{ija} = \sum_j C_{ija} \quad (2)$$

In practice, however, complete convergence is not necessary, as the results do not fundamentally change after 10–20 iterations (Ballas et al., 2005; Anderson, 2007a, as cited in Anderson, 2007b), so that $t = 9$ (for 10 iterations) in this case.

In R, this is implemented with the ipfp package, with `weights_n` as the weights matrix to be filled in each of the iterations, `cons_n` as the total counts from the census, and `ind_cat_n` as the survey data split up as a design matrix (where each individual is a row, each category of the constraints is a column, and the table consists of 0 and 1 for the characteristics that apply to the individual).

From 06_form_sm.R:

```
# Create weights matrix
weights_n <- matrix(data=NA, nrow=nrow(ind), ncol=nrow(cons_n))
dim(weights_n) # check dimensions

# Iterative proportional fitting procedure
library(ipfp)
cons_n <- apply(cons_n,2,as.numeric)
ind_cat_n <- t(ind_cat_n)
x0 <- rep(1,nrow(ind))
weights_n <- as.data.frame(apply(cons_n,MARGIN=1,FUN=function(x)
             ipfp(x,ind_cat_n,x0,maxit=10)))
```

The location constraint has been implemented in a similar way. If the area could not be replicated with the individuals of the same kind of area (see narrow definition in Appendix V, p. 305), then the process was repeated for these areas with the wide location constraint (and again for all of England if necessary):

From 06_form_sm.R:

```
# Checking if individuals could be found with narrow constraint
if (sum(is.nan(as.matrix(weights_n)))>0) {
  nantest <- as.data.frame(1:ncol(weights_n))
  nantest[,"isnanresult"] <- as.data.frame(is.nan(colSums(weights_n[,])))
  nantest <- cbind(nantest,constraints[,1])
```

```
  nantest <- nantest[nantest$isnanresult==TRUE,]
  names(nantest) <- c("row_no","isnanresult","oa11")

  # Same steps as in exhibit above, but with _w instead of _n
  # (...)

  # Combine weights from narrow and wide approach
  for(i in 1:nrow(nantest)){
    weights_n[,nantest$row_no[i]] <- weights_w[,i]
  }
}
```

In order to obtain whole individuals, the individuals were then integerised with the TRS algorithm (Lovelace and Ballas, 2013). This algorithm first of all truncates all weights to the right of the comma, and then replicates the numbers and samples from these. As a consequence, it combines the advantages of probabilistic selection with ensuring that an individual with a weight $w_{ija} \geq 1$ is selected at least once in the respective area. In R, this is implemented as follows:

From 06_form_sm.R:

```
# Load TRS (truncate, replicate, sample)
int_trs <- function(x){
  truncated <- which(x>=1)
  replicated <-rep(truncated,floor(x[truncated]))
  r <- x-floor(x)
  def <- round(sum(x)) - length(replicated)
  if(def==0){
    out <- replicated
  } else {
    out <- c(replicated,sample(length(x),size=def,prob=r,replace=FALSE))
  }
  out
}

# Integerise weights with TRS
set.seed(68649) # Set seed to replicate (run 1)
ints <- unlist(apply(weights_n,2,int_trs))
ints.df <- data.frame(id=ints,zone=(zones_no*prev_no)+
            rep(1:nrow(cons_n),constraints$total))
```

The result is a table with combinations of individuals/areas in an $n : m$ relation. If for a given area a there are 246 individuals, as in the example in Figure 7 (p. 74), then the resulting data frame `ints.df` has 246 rows for this area, potentially with duplicates as some individuals are replicated into this area more than once. Likewise, individuals are replicated into more than one area.

Still missing is the addition of the other values for these individuals as well as the addition of the value for perceived health and health service use based on the lookup matrices created earlier (p. 318). This is done using the following R code:

From 07_form_dataadd.R:

```
# Append health rating category for all three lookup matrices
final.df$healthmain <-
    paste("health",final.df$gender2,final.df$disab2,final.df$age7,sep=".")
final.df$healthnssec <- paste("health",final.df$nssec4,sep=".")
final.df$healtheduc <- paste("health",substr(final.df$ageedu5,0,2),sep=".")
```

```
# Create design matrices for all three lookup matrices
health_mtx_main <- model.matrix(~ final.df$healthmain-1)
health_mtx_nssec <- model.matrix(~ final.df$healthnssec-1)
health_mtx_educ <- model.matrix(~ final.df$healtheduc-1)
colnames(health_matrix_main) # check result

# Multiply lookup matrices to select right value for individual
final.df$health_ratingmain <- round(rowSums(health_mtx_main*
    as.matrix(final.df[,c(91:118)])),2)
final.df$health_ratingnssec <- round(rowSums(health_mtx_nssec*
    as.matrix(final.df[,c(119:123)])),2)
final.df$health_ratingeduc <- round(rowSums(health_mtx_educ*
    as.matrix(final.df[,c(124:128)])),2)

# Take average of all three lookup tables
final.df$health_rating <- rowMeans(final.df[,132:134],na.rm=TRUE)
final.df[1:10,129:135] # Check results
final.df$healthcat <- NULL # Remove health category variable
```

Finally, *08_form_comb.R* combines all datasets into a large `ff` object, as the overall datasets were split into 10 parts to run *06_form_sm.R* and *07_form_dataadd.R* separately for memory reasons with large design matrices and merge commands. Using an `ff` object causes the resulting dataset to be stored on the hard drive rather than in random-access memory (RAM), thereby reducing speed but reducing the risk of running out of RAM for analyses with several million rows.

Validation

For validation, let P_{ija} be the number of individuals with characteristics i and j (as the two example constraint variables) in area a, as obtained based on the weights w_{ija} in the steps outlined above. Furthermore, the total population is denoted by $N = \sum_a C_a = 42{,}989{,}620$ as the sum of all census constraints across all areas, $a_{max} = 171,372$ as the number of areas, and i_{max} and j_{max} as number of categories of the chosen constraint variables.

The main data frames used in validation are called `act` with the actual totals C_{ija} for all categories of all census variables i and j across all areas a, and `sim` with all values P_{ija} aggregated in the same way.

All measures of goodness-of-fit were calculated for each constraint category, and were then reported summarised per constraint and as a total for the simulation (see Appendix VII, p. 308). **Pearson's R** was calculated based on the covariance divided by the product of the variance of each vector and implemented in R as follows:

From 09_valid_int.R:

```
# Compile Pearson's R
pcor <- matrix(data=NA, nrow=ncol(act)-1, ncol=1)
for (i in 2:ncol(act)){
  pcor[i-1,] <- cor(sim[,i],act[,i],method="pearson")
  if (i %in% c(28,29)) {
    pcor[i-1,] <- cor(sim_disab[,i],act_disab[,i],method="pearson")
  }
}
```

The version `act_disab` and `sim_disab` exist due to an error in the census source files (where long-term disability + no long-term disability did not match the total population in the census for 14% of the OAs). While this was handled correctly in the simulation, these areas needed to be discarded in the evaluation to avoid overstated errors.

The **chi-square test** is calculated by squaring the actual and simulated differences and dividing it by the actuals, and then comparing the resulting χ^2 to the theoretical distribution based on degrees of freedom and desired significance level:

$$\chi^2 = \frac{(C_{ija} - P_{ija})^2}{C_{ija}} \tag{3}$$

From 09_valid_int.R:

```
chisq <- matrix(data=NA, nrow=ncol(act)-1, ncol=3)
for (i in 2:ncol(act)){
  chisqdata <- as.data.frame(cbind(act[,i],sim[i]))
  if (i %in% c(28,29)) {
    chisqdata <- as.data.frame(cbind(act_disab[,i],sim_disab[i]))
  }
  chisqdata <- chisqdata[chisqdata[,1]>=5 & chisqdata[,2]>=5,] # exclude cells < 5
  test <- chisq.test(chisqdata) #,simulate.p.value=TRUE)
  chisq[i-1,1] <- get("p.value", test)
  chisq[i-1,2] <- get("statistic", test)
  chisq[i-1,3] <- get("parameter", test)
}
```

The **TAE** is a simple measure that sums up the absolute differences between simulated and actual values and is therefore sensitive to the size of the simulated area:

$$\text{TAE} = \sum_{i,j,a} |C_{ija} - P_{ija}| \tag{4}$$

From 09_valid_int.R:

```
# TAE function
tae_calc <- function(error){
  abs(error)
}

# Compute TAE for age
diff_age<-tae_calc(act[35:41]-sim[35:41])
tae_age <- sum(diff_age)
```

The **SAE** devides the TAE by the population size and thereby makes it more comparable across simulations:

$$\text{SAE} = \frac{\text{TAE}}{N} = \frac{\sum_{i,j,a} |C_{ija} - P_{ija}|}{N} \tag{5}$$

From 09_valid_int.R:

```
# Compute SAE for age
sae_age <- tae_age/sum(act[35:41])
```

The **RMSE** gives more weight to larger deviations and is calculated and implemented as follows:

$$\text{RMSE} = \sqrt{\frac{\sum\limits_{i,j,a} (C_{ija} - P_{ija})^2}{N}} \tag{6}$$

From 09_valid_int.R:

```
# RMSE function
rmse_calc <- function(error){
  sqrt(mean(error^2))
}

# Compute RMSE for age
rmse_age <- rmse_calc(act[35:41]-sim[35:41])
```

Finally, the **NRMSE** further normalises the RMSE to the range of all values:

$$\text{NRMSE} = \frac{\text{RMSE}}{\min(\forall_{ija} C_{ija}) - \max(\forall_{ija} C_{ija})} \tag{7}$$

From 09_valid_int.R:

```
# Compute NRMSE for age
nrmse_age <- rmse_calc(act[35:41]-sim[35:41])/(max(act[35:41])-min(act[35:41]))
```

Deviations larger than 5 or 5% were calculated as the sum of all cases in which the deviation between simulated and observed values was larger than 5 or 5% of the actual values:

$$d_{ija}^{\text{rel}} = \begin{cases} 1, |C_{ija} - P_{ija}| \geq 0.05 \cdot C_{ija} \\ 0, |C_{ija} - P_{ija}| < 0.05 \cdot C_{ija} \end{cases} \tag{8}$$

$$d_a^{\text{rel}} = \frac{\sum\limits_{i,j} d_{ija}^{\text{rel}}}{i_{max} + j_{max}}$$

$$D^{\text{rel}} = \frac{\sum\limits_{a} d_a^{\text{rel}}}{a_{max}}$$

$$d_{ija}^{\text{abs}} = \begin{cases} 1, |C_{ija} - P_{ija}| \geq 5 \\ 0, |C_{ija} - P_{ija}| < 5 \end{cases} \tag{9}$$

$$d_a^{\text{abs}} = \frac{\sum_{i,j} d_{ija}^{\text{abs}}}{i_{max} + j_{max}}$$

$$D^{\text{abs}} = \frac{\sum_{a} d_a^{\text{abs}}}{a_{max}}$$

From 09_valid_int.R:

```
# Relative deviations > 5%
perc5 <- matrix(data=NA, nrow=ncol(act)-1, ncol=1)
for (i in 2:ncol(act)){
  perc5[i-1,] <- length(which(abs(act[i]-sim[i]) > act[i]*0.05))/nrow(act)
  if (i %in% c(28,29)) {
    perc5[i-1,] <- length(which(abs(act_disab[i]-sim_disab[i]) >
        act_disab[i]*0.05))/nrow(act_disab)
  }
}

# Absolute deviations > 5
abs5 <- matrix(data=NA, nrow=ncol(act)-1, ncol=1)
for (i in 2:ncol(act)){
  abs5[i-1,] <- length(which(abs(act[i]-sim[i]) > 5))/nrow(act)
  if (i %in% c(28,29)) {
    abs5[i-1,] <- length(which(abs(act_disab[i]-sim_disab[i]) > 5))/nrow(act_disab)
  }
}
```

For external validation in *10_valid_int.R*, the OLS (2013) and UKHLS (2013) datasets are loaded, recoded, and analysed using cross-tabulations. AMU validation only uses the numbers available in the Ofcom (2013) report.

Analysis

The data was analysed with regressions, SEM and cluster analysis. **Regressions** were run using `biglm` as the main package for the majority of regressions, `nnet` for the multinomial regression in Chapter 5, and `nlme` for the multi-level models in Chapter 6. Below is an example of the analysis in Table 11 (p. 107). All variables were standardised using the `scale` command before.

From 11_analysis_rq1.R:

```
# Linear regression on user access concept (Table 11, Chapter 4)
lr1a <- biglm(health_rating_t ~ gen_nguser + gen_fguser + gen_exuser + num_age +
        num_gender + num_edu + num_nssec + num_disab, data=main.ff)
summary(lm1a)
```

For those analyses that were not available in a "big" version such as `biglm` (which allows for taking the calculation out of RAM), the analysis was done based on a randomly selected subset of 10% of the data (Connelly et al., 2016), averaging the results for 10 random subsamples. In the case above, standard deviations for all standardised coefficients were very small ($0.0002 \leq \sigma \leq 0.0005$), thereby only affecting the third decimal place in a few cases due to rounding differences.

SEM were run in lavaan, using logistic regressions where necessary and the weighted least squares (WLS) estimator for categorical data (Muthén, 1984). With the categorical variables having been declared as ordered before, below is the code that replicates Figure 9 (p. 108).

From 11_analysis_rq1.R:

```
# Structural equation model for user access concept (Figure 9, Chapter 4)

HS.model <- '# regressions
hr_trans ~ gen_nguser + gen_fguser + gen_exuser + num_age + num_gender +
           num_edu + num_nssec + num_disab
gen_nguser ~ med_disabnet + num_age + num_gender +
             num_edu + num_nssec + num_disab
gen_fguser ~ med_disabnet + num_age + num_gender +
             num_edu + num_nssec + num_disab
gen_exuser ~ med_disabnet + num_age + num_gender + num_edu +
             num_nssec + num_disab
med_disabnet ~ hr_trans + num_age + num_gender + num_edu + num_nssec + num_disab
# covariances
gen_nguser ~~ gen_fguser
gen_exuser ~~ gen_fguser
gen_nguser ~~ gen_exuser'

fit2 <- lavaan(HS.model, data=main.ff, auto.var=TRUE, model.type="sem"
    ,int.ov.free=TRUE, estimator="WLSMV", link="logit")
summary(fit2, fit.measures=TRUE)
inspect(fit2,"rsquare")
```

The first step in the **cluster analysis** was calculating the Euclidean distances between each individual for all 8 clustering dimensions. Ward's (1963) method uses squared Euclidean distances, thereby giving greater weight to larger distances:

$$d^2(U,V) = (U_1 - V_1)^2 + (U_2 - V_2)^2 + ... + (U_8 - V_8)^2 \tag{10}$$

Based on these distances $d^2(U,V)$, Ward's methods then calculates the error of the sum of squares (ESS), starting with every individual as its own cluster (ESS = 0) and gradually clusters similar individuals minimising the within cluster sum of squares, so that the increase in the total error of sum of squares ($\triangle \text{ESS}_m$) is as small as possible (Everitt et al., 2011):

$$\text{ESS}_m = \sum_{i=1}^{n} x_{im}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_{im}\right)^2 \tag{11}$$

$$\text{ESS} = \text{ESS}_1 + \text{ESS}_2 + ... + \text{ESS}_m + ... + \text{ESS}_{2150} \tag{12}$$

As a consequence, at the first stage the two individuals with the lowest squared Euclidean distance $d^2(U,V)$ will be merged (Gan et al., 2007). This is repeated until all individuals are in one cluster (see dendrogram, Figure 12). As discussed in the text, the number of clusters is then determined based on the dendrogram, evaluating where there is space before the clusters are reunited at the next level. The snippet below generated the clusters based on the survey data:

From 11_analysis_rq1.R:

```
# Listwise deletion of clustering dimensions
cluster.df <- na.omit(ind[, c("serialno","att_tech","att_learn",
           "per_selfeff","med_trdocs","use_infor","use_interest",
           "use_enjoy","ease_ability","ease_findinfo")])

# Distance calculation (Euclidian distance)
d <- dist(cluster.df[, 2:ncol(cluster.df)], method="euclidean")

# Clustering (Ward's algorithm with squared Euclidean distances)
fit <- hclust(d, method="ward.D2")
groups <- cutree(fit, k=6)
plot(fit)
rect.hclust(fit, k=6, border="grey")

# Addition of other data
serialmerge <- as.data.frame(cbind(cluster.df[,1],groups))
names(serialmerge)[1] <- c("serialno")
ind <- merge(ind[, 1:48], serialmerge, all=TRUE)
```

OAs were mapped to local authority districts with *OA11_WD13_LAD13 _EW_LU* (Census, 2011n), the source files for the maps in Chapter 6 were created as follows:

From 13_analysis_rq3.R:

```
# Create source file with coefficients
lad.df <- coef(lmList(hr_trans2 ~ med_efheal2 + num_age2 + num_gender2 + num_edu2 +
    num_nssec2 + num_disab2 | oalad, data=main.ff, na.action=na.omit))
lad.df <- cbind(Row.Names=rownames(lad.df), lad.df)
rownames(lad.df) <- NULL
lad <- lad[,c(1,3)]
colnames(lad.df) <- c("LAD13CD","med_efheal")
```

Visualisation

These coefficients were visualised subsequently by taking the results tables with the content for each local authority district (see above), and using the shapefiles provided by the ONS (2013a) to create a map. Extreme values were cut (see below) for better visibility. The code below replicates the left map in Figure 24:

From 14_vis_shapefiles.R:

```
# Loading shapefile for Great Britain, selecting England
shapefile <- readShapeSpatial("mapslad/LAD_DEC_2013_GB_BGC.shp")
shapefile_eng <- shapefile[substr(shapefile$LAD13CD, 0, 1)=="E",]

# Mapping the coefficients: Health information finding
healthinfo.df <- merge(shapefile_eng, lad.df, by="LAD13CD")
max(healthinfo.df$med_efheal, na.rm=TRUE)
min(healthinfo.df$med_efheal, na.rm=TRUE)
healthinfo.df$colour <- (rescale(test$med_efheal, to=c(-10,110))*-1)+100
healthinfo.df$colour[healthinfo.df$colour>90] <- 90 # Cut extreme values for visibility
healthinfo.df$colour[healthinfo.df$colour<10] <- 10
healthinfo.df$colour <- trunc(healthinfo.df$colour) # Make whole numbers
hist(healthinfo.df$colour)
healthinfo.df$colour <- paste("gray", healthinfo.df$colour, sep="") # Turn into colour

# Plot the values as a map
plot(shapefile, border="gray88", lwd=0.1)
plot(healthinfo.df, border="gray88", col=healthinfo.df$colour, lwd=0.01, add=TRUE)
```

X Regression diagnostics

Regression assumptions were checked for all regressions in the following way:

1. **Homoscedasticity:** Independence and normality of residuals
2. **Linearity:** Linear relationships of all independent to the dependent variable
3. **Outliers:** No influential points that distort the regression line
4. **Collinearity:** Ideally no relationships between the independent variables

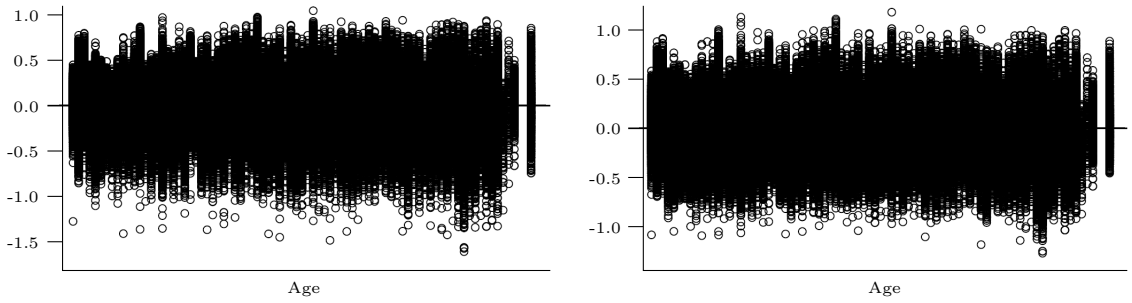
Below the full process for the first regression in this research (Table 11) is laid out. For all other tables, the process was carried out in a similar fashion, so that the results are only summarised on the next page.

Homoscedasticity

Regression residuals were plotted against each of the independent variables to test for **independence of residuals**. In addition, a Q/Q plot was used to test for **normality of residuals**. As a result, perceived health was transformed with a square transformation to the power of 3, and health service use with a square root transformation to the power of 1/3 for all analyses in this research.
Further transformations to the independent variables were only applied if necessary (for example, for non-linearity, see below).

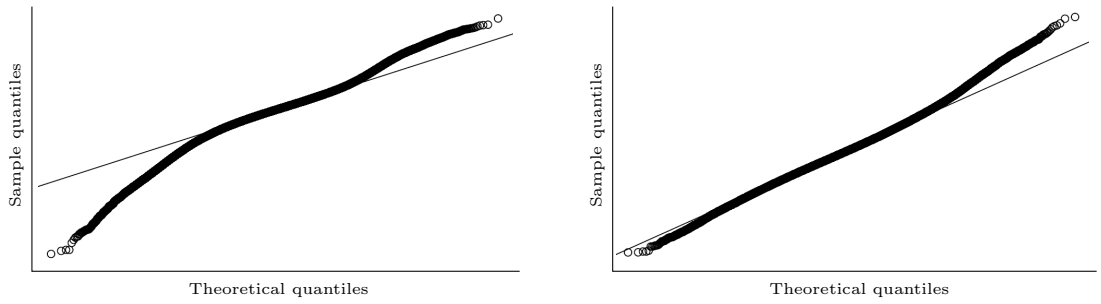
Residuals of regression on perceived health against age

Original *After power transformation*



Q/Q plot of regression on perceived health against age

Original *After power transformation*

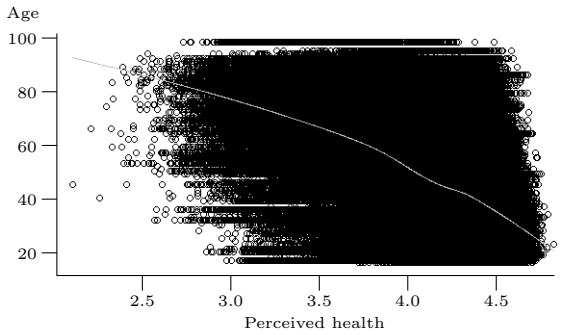


Linearity/outliers

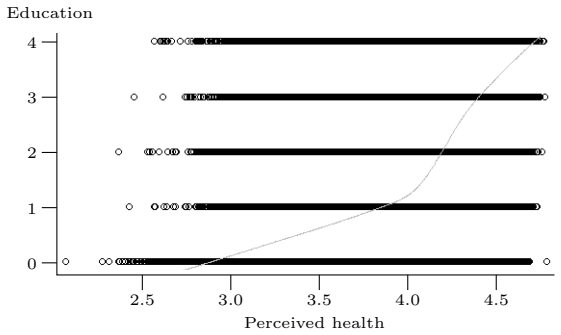
Each of the relationships between the independent and dependent variables was plotted in order to ensure normality using a LOWESS (locally weighted scatterplot smoothing) plot.

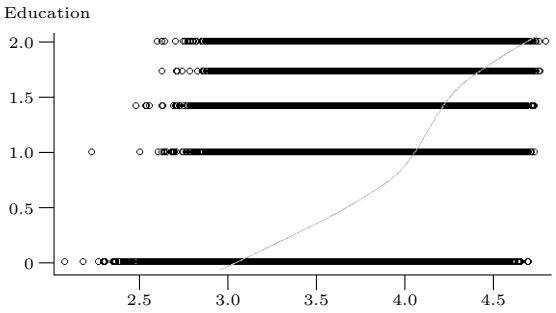
LOWESS of regression on perceived health against age

Original *After power transformation*



[not transformed]





Collinearity

The largest CI was obtained for each regression, and is reported below every analysis in the main document. In all cases, the largest CI was below 15 as the diagnostic boundary (Belsley et al., 1980), indicating that there was no collinearity problem that needed to be fixed. In addition, the highest variance inflation factor (VIF) is reported.

Independent variables	**Condition index: variance decomposition proportions**	**Variance inflation factor**
Next-generation user	0.809	3.576
First-generation user	0.460	2.344
Ex-user	0.095	1.175
Age	0.059	1.817
Gender	0.000	1.013
Education	0.612	1.516
NS-SeC	0.466	1.375
Long-term condition	0.024	1.200
	Largest CI: 6.214	**Largest VIF: 3.576**

These steps were followed for each regression (and SEMs building on them), except for Table 14 and Figure 18, as the regression assumptions can be relaxed for (multinomial) logistic regression. Below are the diagnostics for all regressions:

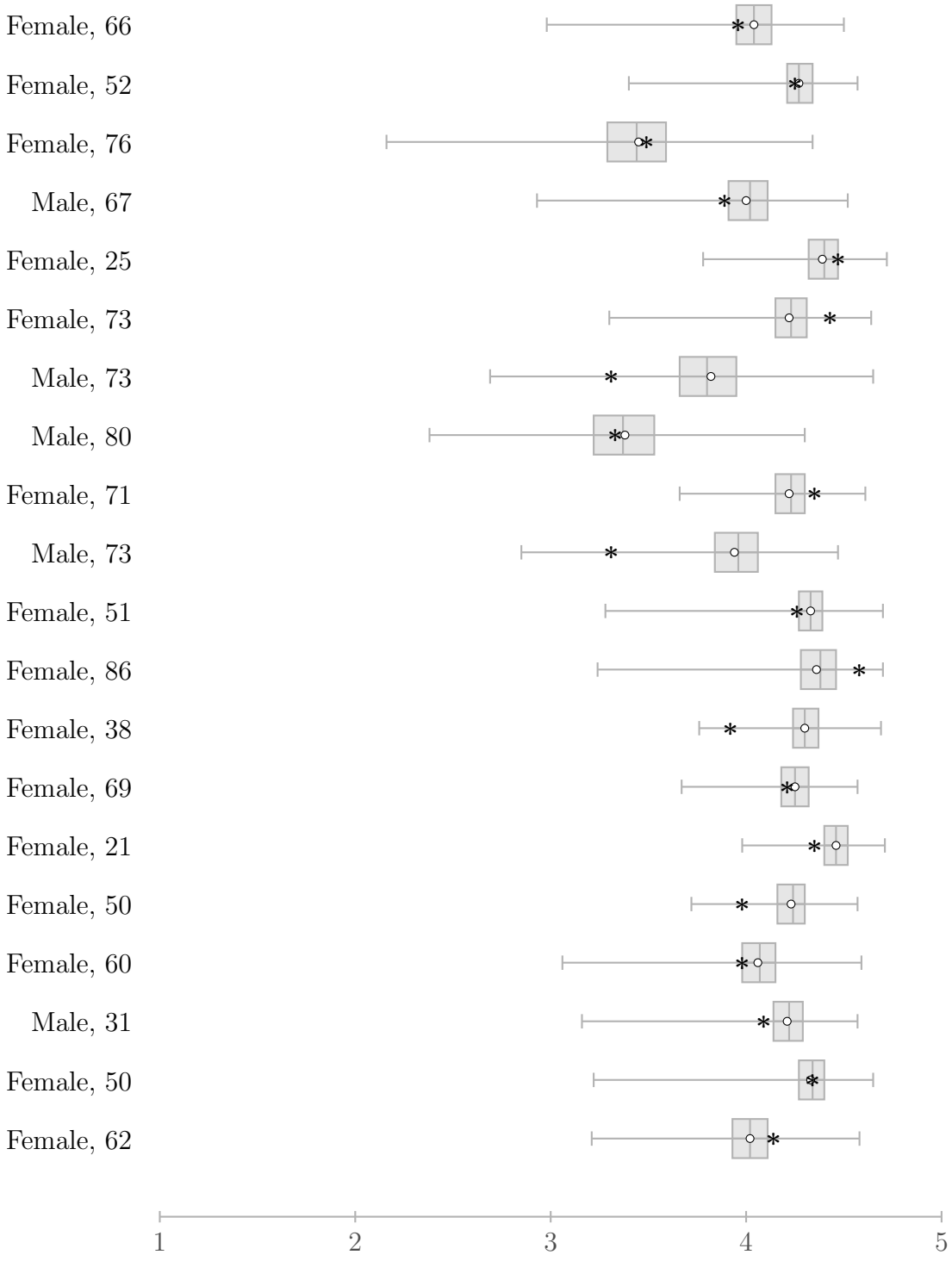
Outcome	Homoscedasticity	Linearity	Outliers	Collinearity
Chapter 4: General regression use/non-use (not reported in table)				
Perceived health	Fulfilled	Education $\sqrt{x}$	–	CI: 5.79
Health service use	Fulfilled	–	–	VIF: 1.67 (user)
Chapter 4: Access divide (Table 11)				
Perceived health	Fulfilled	Education $\sqrt{x}$	–	CI: 6.23
Health service use	Fulfilled	–	–	VIF: 3.58 (next-gen.)
Chapter 4: Support divide (Table 12)				
Perceived health	Fulfilled	Education $\sqrt{x}$	–	CI: 7.63
Health service use	Fulfilled	–	–	VIF: 5.33 (ind. user)
Chapter 4: Usage divide (not reported in table)				
Perceived health	Fulfilled	–	–	CI: 5.37
Health service use	Fulfilled	–	–	VIF: 1.30 (NS-SeC)
Chapter 4: Usage divide (Table 15)				
Perceived health	Fulfilled	–	–	CI: 6.16
Health service use	Fulfilled	–	–	VIF: 3.00 (Learner/u.)
Chapter 6: Age (Table 16)				
Perceived health	Fulfilled	Information importance non-linear, but not transformable	–	CI: 1.77
	Fulfilled			CI: 1.91
	Fulfilled			CI: 2.36
Health service use	Fulfilled	–	–	VIF: 1.32 (NS-SeC)
	Fulfilled			VIF: 1.34 (NS-SeC)
	Fulfilled			VIF: 1.35 (NS-SeC)
Chapter 6: Education (Table 17)				
Perceived health	Fulfilled	–	–	CI: 1.77
	Fulfilled			CI: 1.85
	Fulfilled			CI: 2.04
Health service use	Fulfilled	–	–	VIF: 1.32 (NS-SeC)
	Fulfilled			VIF: 1.35 (NS-SeC)
	Fulfilled			VIF: 1.37 (NS-SeC)
Chapter 6: NS-SeC (Table 18)				
Perceived health	Fulfilled	–	–	CI: 1.77
	Fulfilled			CI: 1.84
	Fulfilled			CI: 2.00

Outcome		Homoscedasticity	Linearity	Outliers	Collinearity
Health service use *continued from Table 18*		Fulfilled	–	–	VIF: 1.33 (NS-SeC)
		Fulfilled			VIF: 1.34 (NS-SeC)
		Fulfilled			VIF: 1.38 (NS-SeC)
Chapter 6: Long-term health condition (Table 19)					
Perceived health		Fulfilled	–	–	CI: 1.77
		Fulfilled			CI: 1.77
		Fulfilled			CI: 1.81
Health service use		Fulfilled	–	–	VIF: 1.33 (NS-SeC)
		Fulfilled			VIF: 1.33 (NS-SeC)
		Fulfilled			VIF: 1.34 (NS-SeC)
Chapter 6: Social environment (Table 20)					
Perceived health		Fulfilled	–	–	CI: 1.77
		Fulfilled			CI: 1.89
		Fulfilled			CI: 1.95
Health service use		Fulfilled	–	–	VIF: 1.33 (NS-SeC)
		Fulfilled			VIF: 1.38 (NS-SeC)
		Fulfilled			VIF: 1.40 (NS-SeC)
Chapter 6: Perceived health across regions (Table 21)					
Perc. health	*London*	Fulfilled	–	–	CI: 1.87; VIF: 1.31
	South East	Fulfilled	–	–	CI: 1.82; VIF: 1.33
	South West	Fulfilled	–	–	CI: 1.84; VIF: 1.34
	East of E.	Fulfilled	–	–	CI: 1.81; VIF: 1.33
	East Mid.	Fulfilled	–	–	CI: 1.81; VIF: 1.36
	West Mid.	Fulfilled	–	–	CI: 1.86; VIF: 1.38
	Yorkshire	Fulfilled	–	–	CI: 1.84; VIF: 1.33
	North East	Fulfilled	–	–	CI: 1.86; VIF: 1.34
	North West	Fulfilled	–	–	CI: 1.81; VIF: 1.33
Chapter 6: Perceived health across urban/rural areas (Table 22)					
Perc. health	*Conurbation*	Fulfilled	–	–	CI: 1.83; VIF: 1.33
	City/town	Fulfilled	–	–	CI: 1.79; VIF: 1.33
	Town/fringe	Fulfilled	–	–	CI: 1.90; VIF: 1.33
	Village	Fulfilled	–	–	CI: 2.21; VIF: 1.38
	Hamlet	Fulfilled	–	–	CI: 2.28; VIF: 1.40
Chapter 6: Health service use across urban/rural areas (Table 23)					
Health service use	*Conurbation*	Fulfilled	–	–	CI: 1.82; VIF: 1.35
	City/town	Fulfilled	–	–	CI: 1.80; VIF: 1.35
	Town/fringe	Fulfilled	–	–	CI: 1.87; VIF: 1.35
	Village	Fulfilled	–	–	CI: 2.15; VIF: 1.40
	Hamlet	Fulfilled	–	–	CI: 2.21; VIF: 1.42

XI Results for OxIS versus simulated dataset

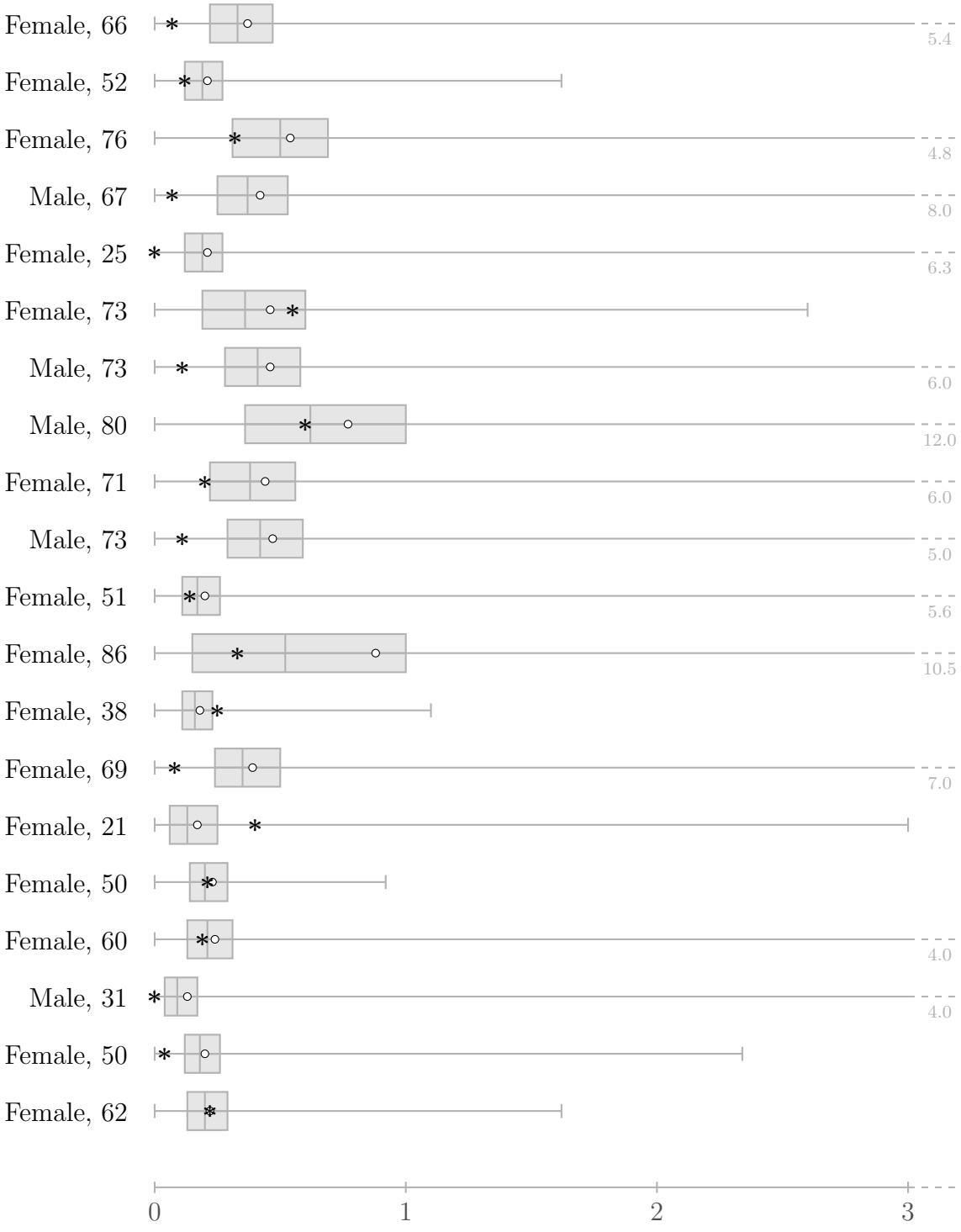
Perceived health: spread for SM model versus OxIS individuals

Below is the spread of the simulated outcome variable of perceived health for 20 randomly selected individuals (median = |, mean = ∘, hinges denote the range). The value of the OxIS individual in his/her own OAs is provided as a star (*).



Health service use: spread for SM model versus OxIS individuals

Below is the spread of the simulated outcome variable of health service use for a specific individual based on the SM dataset (median = |, mean = •) for the same 20 individuals. The value of the OxIS individual in his/her own OAs is provided as a light dot (○).



Comparison of OxIS versus simulated data results for selected analyses

The tables below replicate all regressions based on OxIS data with perceived health and health service use appended directly. The regressions in section 6.3 were not replicated due to the low N for OxIS when using subsets. Significance codes for OxIS are *** < 0.001; ** < 0.01 and * < 0.05 (not provided for SM, as all coefficients are significant). As explained in subsection 3.4.3, Bonferroni correction has not been applied, as there are mathematical, logical and practical challenges associated to doing so (Moran, 2003; Garamszegi, 2006; Cohen, 1990).

Regression based on access divide (Table 11)

	Perceived health		**Health service use**	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
Next-generation user	0.067	0.061 **	−0.037	−0.035 *
First-generation user	0.034	0.020	−0.017	−0.038 *
Ex-user	0.004	0.023	0.015	−0.026
Never-user *(omitted)*	–	–	–	–
Age	−0.207	−0.159 ***	0.433	0.381 ***
Gender	0.094	0.126 ***	0.150	0.098 ***
Education	0.314	0.304 ***	−0.031	−0.016
NS-SeC	−0.083	−0.011	0.033	0.014
Long-term health condition	−0.462	−0.489 ***	0.030	0.032 *
Adjusted R^2	**0.60**	**0.55**	**0.24**	**0.18**
N	**42,989,620**	**2,150**	**42,989,620**	**2,150**
Largest CI	**6**	**3**	**6**	**3**

Regression based on support divide (Table 12)

	Perceived health		**Health service use**	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
Independent user	0.071	0.094 **	−0.061	−0.139 ***
Supported user	0.045	0.080 ***	−0.030	−0.099 **
Unsupported user	0.026	0.038	−0.023	−0.057 *
Proxy access	0.016	0.040 *	−0.007	−0.054 *
Proxy availability	0.009	0.072 ***	−0.010	−0.010
Fully excluded *(omitted)*	–	–	–	–
Age	−0.218	−0.180 ***	0.437	0.375 ***
Gender	0.092	0.124 ***	0.150	0.099 ***
Education	0.316	0.307 ***	−0.031	−0.010
NS-SeC	−0.084	−0.014	0.032	0.009
Long-term health condition	−0.462	−0.482 ***	0.030	0.028
Adjusted R^2	**0.60**	**0.56**	**0.24**	**0.18**
N	**42,989,620**	**2,150**	**42,980,620**	**2,150**
Largest CI	**8**	**5**	**8**	**5**

Regression based on usage divide (see section 4.2.3, not shown in table)

	Perceived health		Health service use	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
Health infomation seeker	−0.005	−0.002	0.009	0.005
No health info. seeker	–	–	–	–
Age	−0.206	−0.162 ***	0.414	0.360 ***
Gender	0.094	0.113 ***	0.202	0.133 ***
Education	0.263	0.296 ***	−0.005	−0.001
NS-SeC	−0.134	−0.008	0.055	0.027
Long-term health condition	−0.461	−0.533 ***	0.032	0.018
Adjusted R^2	**0.46**	**0.48**	**0.20**	**0.15**
N	**42,989,620**	**1,513**	**42,989,620**	**1,513**
Largest CI	**2**	**2**	**2**	**2**

Regression with typology of health information seekers (Table 15)

	Perceived health		Health service use	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
User of learner type	0.073	0.068 **	−0.043	−0.028
User of pragmatist type	0.047	0.030	−0.031	−0.035
User of sceptic type	0.038	0.019	−0.023	−0.019
User of worrier type	0.041	0.028	−0.011	−0.000
User of delegator type	0.027	0.012	−0.009	−0.007
User of adigital type	0.016	−0.014	−0.010	−0.017
Non-user of learner type	0.014	0.026 *	−0.000	−0.017
Non-user of pragmatist type	−0.005	−0.014	0.009	−0.004
Non-user of sceptic type	0.005	0.025 *	−0.002	−0.047 *
Non-user of worrier type	−0.007	0.028 *	0.017	−0.002
Non-user of delegator type	0.023	0.018	−0.006	0.025
Non-user of adigital type	–	–	–	–
Age	−0.208	−0.161 ***	0.433	0.377 ***
Gender	0.096	0.130 ***	0.148	0.099 ***
Education	0.306	0.306 ***	−0.032	−0.020
NS-SeC	−0.086	−0.012	0.035	0.014
Long-term health condition	−0.473	−0.487 ***	0.032	0.033
Adjusted R^2	**0.60**	**0.56**	**0.24**	**0.18**
N	**42,989,620**	**2,150**	**42,980,620**	**2,150**
Largest CI	**6**	**4**	**6**	**4**

Regression for age effects (Table 16)

	Perceived health		Health service use	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
Improved my health online	0.018	0.005	−0.015	−0.003
× Age	0.008	−0.016	0.008	−0.008
Internet interest	0.016	0.031	−0.019	0.019
× Age	−0.015	−0.054 *	0.025	0.066 *
× Improved health	0.025	−0.017	0.016	0.014
× Improved health × Age	0.013	−0.031	0.001	−0.014
Internet info. importance	0.024	0.030	−0.002	−0.033
× Age	−0.020	−0.042 *	−0.014	−0.041
× Improved health	0.003	−0.020	−0.011	−0.021
× Improved health × Age	0.010	0.027	0.016	0.021
Age	−0.206	−0.098 ***	0.412	0.316 ***
Gender	0.102	0.112 ***	0.200	0.156 ***
Education	0.327	0.293 ***	−0.030	−0.024
NS-SeC	−0.101	−0.005	0.034	0.026
Long-term health condition	−0.476	−0.528 ***	0.034	0.056 *
Adjusted R^2	**0.48**	**0.48**	**0.21**	**0.16**
N	**33,874,471**	**1,511**	**33,874,471**	**1,511**
Largest CI	**2**	**4**	**2**	**4**

Regression for socio-economic status effects (Table 18)

	Perceived health		Health service use	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
Improved my health online	0.019	0.006	−0.019	−0.006
× NS-SeC	0.008	0.016 **	0.006	0.008
Technology attitude	0.044	0.060	−0.014	−0.012
× NS-SeC	0.021	0.007	−0.019	0.005
× Improved health	−0.006	−0.001	−0.004	−0.014
× Improved health × NS-SeC	−0.018	−0.002	−0.004	0.019
Self-efficacy	0.009	−0.004	0.009	0.008
× NS-SeC	0.049	0.041 *	−0.007	−0.019
× Improved health	0.019	0.009	−0.001	0.012
× Improved health × NS-SeC	−0.037	−0.016	0.016	−0.021
Age	−0.193	−0.134 ***	0.416	0.341 ***
Gender	0.101	0.116 ***	0.200	0.156 ***
Education	0.307	0.296 ***	−0.027	−0.020
NS-SeC	−0.108	−0.002	0.040	0.023
Long-term health condition	−0.485	−0.524 ***	0.031	0.048 *
Adjusted R^2	**0.49**	**0.48**	**0.21**	**0.16**
N	**33,896,196**	**1,513**	**33,896,196**	**1,513**
Largest CI	**2**	**3**	**2**	**3**

Regression for education effects (Table 17)

	Perceived health		Health service use	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
Improved my health online	0.020	0.010	−0.019	−0.016
× Education	−0.049	−0.021	0.026	0.036
Trust in doctors	0.007	0.033 *	−0.007	0.012
× Education	−0.035	−0.042 *	−0.000	0.015
× Improved health	0.018	−0.000	−0.023	−0.019
× Improved health × Educ.	0.009	0.013	−0.003	0.024
Information finding ability	0.010	0.023	−0.003	−0.004
× Education	−0.003	0.012	−0.006	−0.034
× Improved health	−0.011	−0.003	0.001	0.003
× Improved health × Educ.	0.007	−0.003	0.006	0.026
Information evaluation ability	0.051	0.021	−0.019	0.008
× Education	0.000	−0.014	−0.013	−0.014
× Improved health	−0.017	−0.012	0.000	−0.029
× Improved health × Educ.	−0.008	−0.021	−0.005	0.002
Age	−0.201	−0.159 ***	0.418	0.343 ***
Gender	0.100	0.109 ***	0.202	0.159 ***
Education	0.309	0.299 ***	−0.025	−0.008
NS-SeC	−0.101	−0.007	0.035	0.029
Long-term health condition	−0.489	−0.528 ***	0.033	0.052 *
Adjusted R^2	**0.49**	**0.48**	**0.21**	**0.16**
N	**33,896,196**	**1,513**	**33,896,196**	**2,150**
Largest CI	**2**	**4**	**2**	**4**

Regression for long-term health condition effects (Table 19)

	Perceived health		Health service use	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
Improved my health online	0.020	0.004	−0.020	−0.013
× Long-term health condition	0.008	0.021	0.019	0.020
Self-efficacy	0.021	0.005	0.004	0.013
× Long-term health condition	0.010	0.000	0.006	0.011
× Improved health	−0.028	−0.001	0.002	0.017
× Improved health × L.-t. con.	−0.014	−0.017	−0.012	0.041
Age	−0.210	−0.151 ***	0.421	0.344 ***
Gender	0.100	0.112 ***	0.202	0.156 ***
Education	0.321	0.305 ***	−0.031	−0.023
NS-SeC	−0.106	−0.003	0.038	0.028
Long-term health condition	0.487	−0.527 ***	0.031	0.053 *
Adjusted R^2	**0.48**	**0.48**	**0.21**	**0.16**
N	**33,896,196**	**2,150**	**33,896,196**	**1,513**
Largest CI	**2**	**2**	**2**	**2**

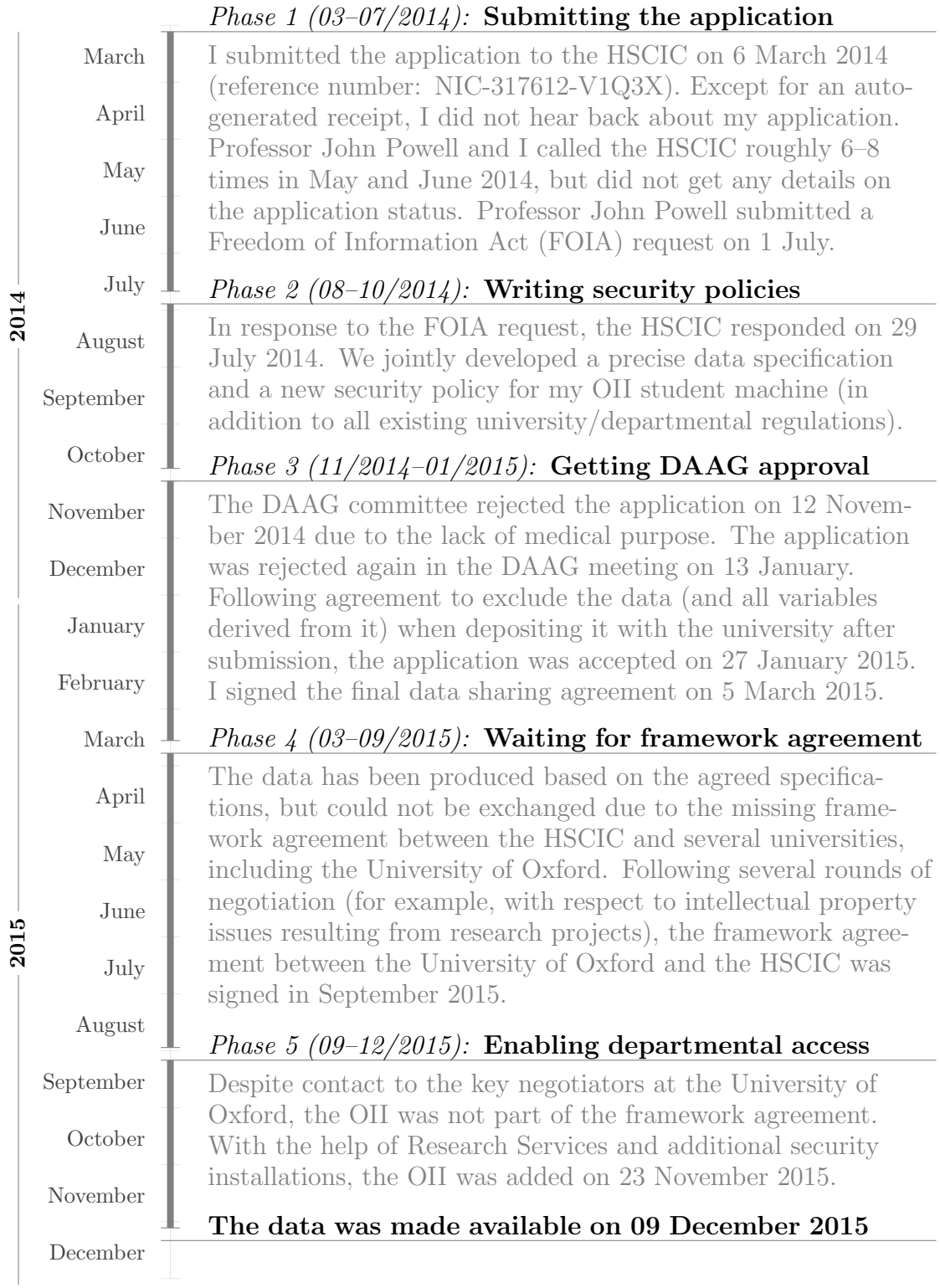
Regression for social environment effects (Table 20)

	Perceived health		Health service use	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
Improved my health online	0.017	0.001	−0.017	−0.008
Help availability	0.028	0.028	−0.013	−0.009
× Improved health	−0.006	0.008	−0.000	0.008
Information evaluation ability	0.049	−0.011	−0.020	−0.001
× Improved health	−0.027	−0.000	0.007	−0.017
× Help availability	−0.014	−0.008	−0.014	−0.040
× Help avail. × Improved health	0.000	−0.023	−0.002	−0.010
Social embeddedness	0.047	0.013	0.016	0.001
× Improved health	0.010	0.012	−0.015	−0.031
Medical expertise in household	−0.025	−0.020	0.004	−0.002
× Improved health	−0.003	−0.007	0.008	−0.004
Age	−0.199	−0.166 ***	0.419	0.351 ***
Gender	0.107	0.111 ***	0.199	0.158 ***
Education	0.325	0.297 ***	−0.026	−0.016
NS-SeC	−0.099	−0.016	0.036	0.031
Long-term health condition	−0.471	−0.533 ***	0.034	0.035
Adjusted R^2	**0.48**	**0.47**	**0.21**	**0.15**
N	**33,896,196**	**1,513**	**33,896,196**	**2,150**
Largest CI	**2**	**3**	**2**	**3**

Baseline tables (first table: all individuals; second table: users only)

	Perceived health		Health service use	
Independent variables	β (SM)	β (OxIS)	β (SM)	β (OxIS)
Age	−0.232	−0.186 ***	0.447	0.392 ***
Gender	0.092	0.125 ***	0.150	0.100 ***
Education	0.328	0.315 ***	−0.040	−0.023
NS-SeC	−0.091	−0.016	0.039	0.016
Long-term health condition	−0.468	−0.489 ***	0.035	0.035
Adjusted R^2	**0.59**	**0.55**	**0.24**	**0.18**
N	**42,989,620**	**2,150**	**42,989,620**	**2,150**
Largest CI	**2**	**2**	**2**	**2**
Age	−0.206	−0.168 ***	0.448	0.363 ***
Gender	0.084	0.130 ***	0.182	0.182 ***
Education	0.305	0.318 ***	−0.031	−0.023
NS-SeC	−0.084	−0.016	0.037	0.040
Long-term health condition	−0.475	−0.500 ***	0.033	0.030
Adjusted R^2	**0.47**	**0.47**	**0.21**	**0.16**
N	**33,896,196**	**1,513**	**33,896,196**	**1,513**
Largest CI	**2**	**2**	**2**	**2**

XII Timeline of HES data request



XIII Area selection

Diversity across area classifications

Area classification	**Districts & 2nd level area classification**
Group 1: *Rural residents*	Wiltshire (1a) Maidstone (1b) Newcastle-under-Lyme (1b) Kings Lynn & West Norfolk (1c)
Group 2: *Cosmopolitans*	Ealing (2d)
Group 3: *Ethnicity central*	Hammersmith & Fulham (3d)
Group 4: *Multicultural metropolitans*	Birmingham (4a)
Group 5: *Urbanites*	Cornwall (5b)
Group 6: *Suburbanites*	East Dorset (6a) Isle of Wight (6b)
Group 7: *Constrained city dwellers*	Gateshead (7d)
Group 8: *Hard-pressed living*	Amber Valley (8a) South Norfolk (8c) Middlesbrough (8d)

Diversity across urban/rural classifications and regions

Urban/rural classifications		**Districts & region**
Urban	*Conurbation*	Ealing (London) Hammersmith & Fulham (London) Birmingham (West Midlands) Gateshead (North East) Amber Valley (East Midlands)
	City/town	Maidstone (South East) East Dorset (South West) Isle of Wight (South East) Middlesbrough (North East) South Norfolk (East)
Rural	*Town/fringe*	King's Lynn & West Norfolk (East) Cornwall (South West)
	Village	Newcastle-under-Lyme (West Midlands)
	Hamlet	Wiltshire (South West)

Four areas were selected from the backup list: Newcastle-under-Lyme, Wiltshire, East Dorset, and South Norfolk. This was necessary as in Northumberland and Suffolk Coastal, no interviewee could be recruited from the former OxIS participants. In Mid Suffolk and Bristol, data quality problems in relation to the OA recorded for the participants required changes to the area selection.

XIV Interviewee health conditions

Below are the conditions of interviewees, someone they cared for (partner, spouse, child) and family/friends based on the MCD classification (European Commission, 2015).

Conditions	Interviewee	Care receiver	Friends/family
Cardiovascular diseases (including hypertension)	– Stroke – Heart failure – Septal defects – Deep vein thrombosis (DVT) – Hypertension	– Heart failure – Hypertension	– Septal defects – Hypertension
Metabolic diseases	– Diabetes – Hypothyroidism	– Diabetes – Irritable Bowel Syndrome (IBS)	
Respiratory diseases	– Asthma	– Asthma	– Chronic obstructive pulmonary disease (COPD)
Musculoskeletal diseases	– Arthritis, gout – Disintegrated spine, spinal fracture – Other back pains	– MS	– Gout
Renal diseases	– Kidney stones		
Cancer		– Skin cancer	– Gastric cancer – Pancreatic cancer
Neurodegenerative diseases			– Parkinson's disease – Dementia
Mental disorders	– Depression – ADHD – OCD	– Depression – ADHD	– Depression
Epilepsy	– Epilepsy		– Epilepsy
Autistic spectrum disorders		– Autism – Asperger's Syndrome	
Visual impairment	– Glaucoma – Cataract – Shortsightedness		
Hearing disorders	– Hearing loss	– Hearing loss	
Other aspects of medical relevance (unclassified)	– Pregnancy/IVF – Men's health – Alcoholism – Smoking – Dieting/weight	– Eczema	– Rare chromosome disorder

XV Overview of interviewees

Interviewee overview		Internet use	Health characteristics
Amber Valley – East Midlands, conurbation, 8a "Industrious communities"			
Ahmad	Education level 3	Next-generation user	Health rating: 4.5
Male, 24	NS-SeC 2	Independent user	Weight loss, sports injuries
Non-OxIS	Racing motorcyclist	Learner type	
Anthony	Education level 4	First-generation user	Health rating: 4 (p. 105)
Male, 58	NS-SeC 1	Independent user	Cataract; wife with knee
OxIS	Engineer	Worrier type	replacement
Birmingham – West Midlands, conurbation, 4a "Rented family living"			
Barbara	Education level 2	Next-generation user	Health rating: 4 (p. 106)
Female, 41	long-term unemployed	Supported user	Depression, alcoholism
OxIS	None / not available	Sceptic type	
Brenda	Education level 2	First-generation user	Health rating: 4
Female, 40	NS-SeC 3	Independent user	Sports injuries; son with
OxIS	Full-time carer for son	Pragmatist type	autism
Brian	Education level 3	Next-generation user	Health rating: 4
Male, 54	NS-SeC 1	Independent user	Eye surgery
Non-OxIS	Middle manager	Sceptic type	
Cornwall – South West, rural town/fringe, 5b "Ageing urban living"			
Carol	Education level 1	First-generation user	Health rating: 4
Female, 61	NS-SeC 2	Independent user	Hypothyroidism; husband
Non-OxIS	Inn owner	Pragmatist type	with heart failure
Catherine	Education level 1	Never-user	Health rating: 4 (p. 106)
Female, 77	NS-SeC 3	Proxy availability	Disintegrated spine
OxIS	Textile worker	Adigital type	
Charlotte	Education level 0	Never-user	Health rating: 4
Female, 78	NS-SeC 2	Proxy access	DVT, family with lung
OxIS	Sales representative	Delegator type	cancer and septic defects
Courtney	Education level 1	Next-generation user	Health rating: 4
Female, 29	NS-SeC 3	Independent user	
OxIS	Cleaner	Pragmatist type	
East Dorset – South West, urban city/town, 6a "Suburban achievers"			
Danielle	Education level 3	Next-generation user	Health rating: 4
Female, 30	NS-SeC 1	Independent user	Son with Eczema and
OxIS	Financial analyst	Learner type	asthma
David	Education level 2	Next-generation user	Health rating: 4
Male, 65	NS-SeC 1	Independent user	Hypertension
OxIS	Aircraft engineer	Pragmatist type	

Interviewee overview		Internet use	Health characteristics
Donald	Education level 4	First-generation user	Health rating: 4
Male, 74	NS-SeC 3	Independent user	Knee replacement
OxIS	Mainframe developer	Sceptic type	
Ealing – London, conurbation, 2d "Aspiring and affluent"			
Elizabeth	Education level 2	Ex-user	Health rating: n/a
Female, 55	long-term unemployed	Proxy access	Stroke
OxIS	Sociology student	Sceptic type	
Ethan	Education level 4	Next-generation user	Health rating: 4
Male, 69	NS-SeC 1	Independent user	
Non-OxIS	Electrical engineer	Sceptic type	
Gateshead – North East, conurbation, 7d "Ageing city dwellers"			
Gary	Education level 0	Never-user	Health rating: n/a
Male, 61	NS-SeC 3	Proxy availability	
Non-OxIS	Cleaner	Adigital type	
George	Education level 1	Ex-user	Health rating: 3
Male, 69	NS-SeC 2	Proxy availability	Gout
Non-OxIS	Police officer (retired)	Adigital type	
Hammersmith and Fulham – London, conurbation, 3d "Aspirational techies"			
Harold	Education level 4	Next-generation user	Health rating: n/a
Male, 39	NS-SeC 1	Independent user	Kidney stones
OxIS	Public house licensee	Learner type	
Helen	Education level 4	Next-generation user	Health rating: 3–4
Female, 43	NS-SeC 1	Independent user	Epilepsy; niece with rare
OxIS	Designer	Worrier type	chromosome disorder
Hugh	Education level 4	Next-generation user	Health rating: 4
Male, 45	NS-SeC 1	Independent user	
OxIS	Head of intl. company	Pragmatist type	
Isle of Wight – South East, urban city/town, 6b "Semi-detached suburbia"			
Ian	Education level 1	Next-generation user	Health rating: 2–5 (p. 106)
Male, 42	Long-term unemployed	Independent user	OCD, men's health, septal
Non-OxIS	Appliances inspector	Learner type	heart defect
Irene	Education level 0	Never-user	Health rating: 4
Female, 82	NS-SeC 3	Proxy access	Arthritis
OxIS	Receptionist	Delegator type	
Irvin	Education level 0	Never-user	Health rating: 4
Male, 85	NS-SeC 1	Proxy access	
Non-OxIS	Airforce officer (retired)	Delegator type	
Isabelle	Education level 2	Next-generation user	Health rating: 4
Female, 38	NS-SeC 3	Independent user	Son with autism
OxIS	Carer	Learner type	

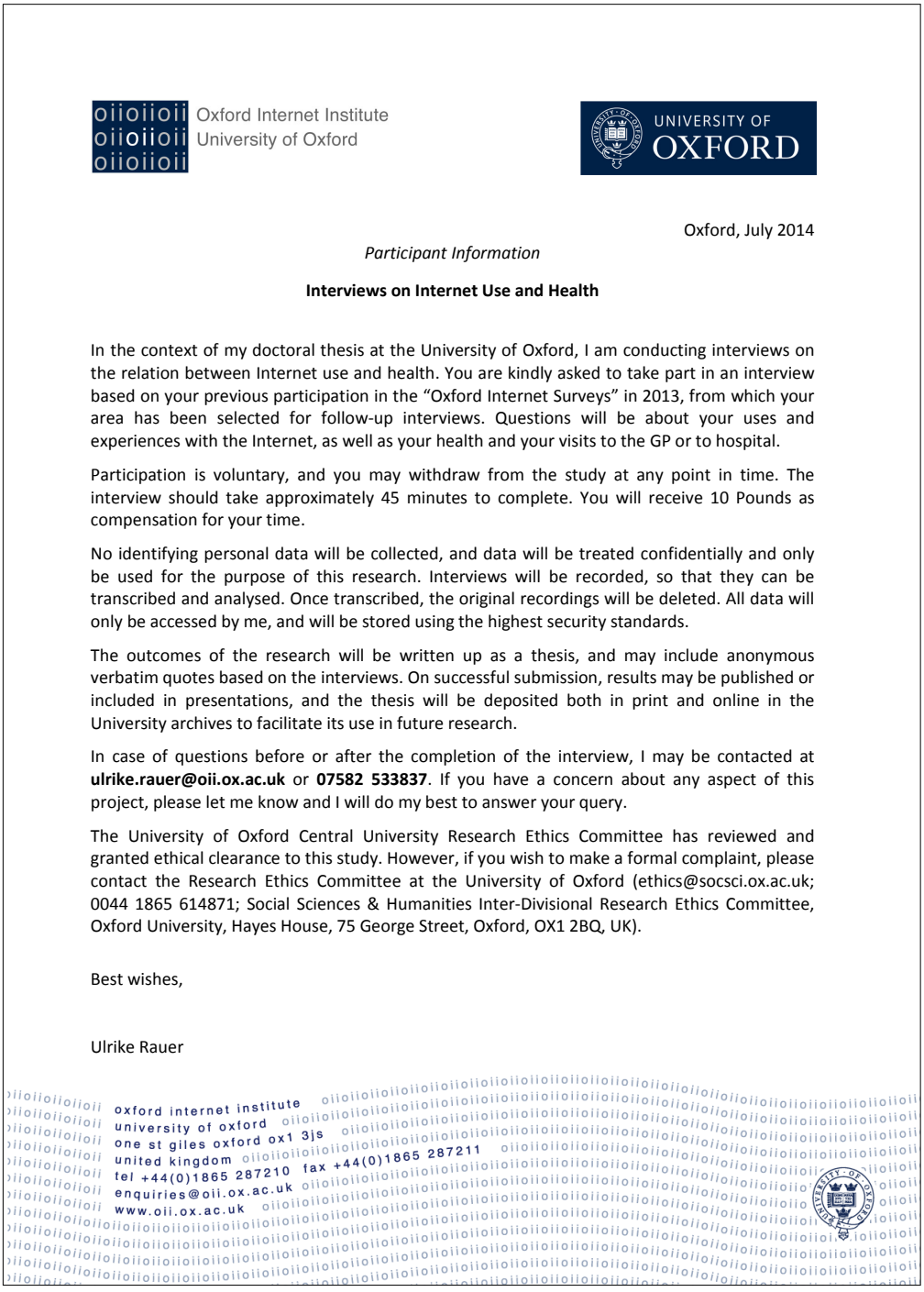
Interviewee overview		Internet use	Health characteristics
King's Lynn & West Norfolk – East, rural town/fringe, 1c "Ageing rural dwellers"			
Karen	Education level 1	First-generation user	Health rating: 5 (p. 106)
Female, 69	NS-SeC 3	Supported user	Glaucoma
OxIS	Cook	Worrier type	
Kathleen	Education level 0	Never-user	Health rating: 2–4 (p. 106)
Female, 75	NS-SeC 1	Proxy access	Back problems, hypertension
OxIS	Civil servant	Delegator type	
Kevin	Education level 1	Never-user	Health rating: 1
Male, 40	long-term unemployed	Fully excluded	Stroke
OxIS	Mechanical engineer	Adigital type	
Maidstone – South East, urban city/town, 5a "Urban professionals/families"			
Margaret	Education level 2	First-generation user	Health rating: 5
Female, 57	NS-SeC 3	Independent user	Back problems
OxIS	School social worker	Pragmatist type	
Martha	Education level 4	Next-generation user	Health rating: 4 (p. 106)
Female, 34	NS-SeC 1	Independent user	Pregnancy/IVF
Non-OxIS	Make-up artist	Pragmatist type	
Mary	Education level 1	Next-generation user	Health rating: 3
Female, 26	NS-SeC 3	Supported user	ADHD, asthma; son with ADHD
OxIS	Nursery worker	Delegator type	
Matthew	Education level 3	Next-generation user	Health rating: 3–3.5
Male, 47	NS-SeC 2	Independent user	Spinal fracture, hypertension
OxIS	Builder (self-employed)	Pragmatist type	
Middlesbrough – North East, urban city/town, 8d "Migration and churn"			
Michael	Education level 4	Next-generation user	Health rating: 5
Male, 22	full-time student	Independent user	
Non-OxIS	Architecture student	Learner type	
Mildred	Education level 0	Never-user	Health rating: n/a
Female, 91	NS-SeC 2	Proxy availability	Hearing problems
OxIS	Shop assistant	Adigital type	
Miriam	Education level 3	Next-generation user	Health rating: 3
Female, 58	NS-SeC 2	Independent user	Stroke, depression, arthritis
Non-OxIS	Nursery worker	Learner type	
Newcastle-under-Lyme – West Midlands, rural village, 1b "Rural tenants"			
Natalie	Education level 1	Never-user	Health rating: 5
Female, 31	NS-SeC 3	Proxy access	Partner with diabetes; son with hearing disorder
OxIS	Administrator	Delegator type	
Nathan	Education level 1	Next-generation user	Health rating: 4
Male, 90	NS-SeC 1	Independent user	Heart failure
OxIS	Company manager (retired)	Pragmatist type	

Interviewee overview		Internet use	Health characteristics
Nicolas	Education level 4	Next-generation user	Health rating: 5
Male, 19	full-time student	Independent user	Sports injuries
OxIS	n/a	Pragmatist type	
Norman	Education level 4	Never-user	Health rating: 5
Male, 58	NS-SeC 1	Proxy access	
OxIS	Clergyman	Delegator type	
South Norfolk – East, urban city/town, 8c "Hard pressed ageing workers"			
Sarah	Education level 1	Next-generation user	Health rating: 3.5–4
Female, 30	NS-SeC 3	Independent user	
OxIS	Delivery driver	Adigital type	
Shirley	Education level 2	First-generation user	Health rating: n/a
Female, 71	NS-SeC 3	Supported user	
Non-OxIS	Profess. carer (retired)	Adigital type	
Susan	Education level 0	First-generation user	Health rating: 4
Female, 58	NS-SeC 3	Supported user	Thyroid problems
OxIS	Cleaner	Adigital type	
Wiltshire – South West, rural hamlet, 1a "Farming communities"			
Wanda	Education level 1	Next-generation user	Health rating: 3–4
Female, 58	NS-SeC 2	Supported user	Arthritis
OxIS	Veterinary assistant	Pragmatist type	
Whitney	Education level 4	First-generation user	Health rating: 4
Female, 60	NS-SeC 1	Independent user	Husband with MS
OxIS	Mgmt. services trainer	Sceptic type	
Wynne	Education level 4	First-generation user	Health rating: 4–5
Female, 65	NS-SeC 1	Supported user	Husband with gout
OxIS	Crop farmer	Worrier type	

Alternative modes of interviews: Telephone interviews were conducted with *Elizabeth* and *Harold*. The interview with *Barbara* took place in two separate sessions based on the interviewee's request.

Recruitment of non-OxIS interviewees: *Ahmad*, *Carol*, *Gary*, *George*, *Martha*, *Michael* and *Miriam* were approached in public spaces within the OA. *Ethan* was referred through a personal contact, as he happened to live in the same OA used in OxIS. *Shirley* was recommended by her neighbour *Susan*. *Brian* was recruited by his fiancée *Barbara*. *Ian* and *Irvin* were suggested by their respective wives *Isabelle* and *Irene*.

XVI Interview participant documents



Oxford Internet Institute
University of Oxford

UNIVERSITY OF OXFORD

Oxford, July 2014

Participant Information

Interviews on Internet Use and Health

In the context of my doctoral thesis at the University of Oxford, I am conducting interviews on the relation between Internet use and health. You are kindly asked to take part in an interview based on your previous participation in the "Oxford Internet Surveys" in 2013, from which your area has been selected for follow-up interviews. Questions will be about your uses and experiences with the Internet, as well as your health and your visits to the GP or to hospital.

Participation is voluntary, and you may withdraw from the study at any point in time. The interview should take approximately 45 minutes to complete. You will receive 10 Pounds as compensation for your time.

No identifying personal data will be collected, and data will be treated confidentially and only be used for the purpose of this research. Interviews will be recorded, so that they can be transcribed and analysed. Once transcribed, the original recordings will be deleted. All data will only be accessed by me, and will be stored using the highest security standards.

The outcomes of the research will be written up as a thesis, and may include anonymous verbatim quotes based on the interviews. On successful submission, results may be published or included in presentations, and the thesis will be deposited both in print and online in the University archives to facilitate its use in future research.

In case of questions before or after the completion of the interview, I may be contacted at **ulrike.rauer@oii.ox.ac.uk** or **07582 533837**. If you have a concern about any aspect of this project, please let me know and I will do my best to answer your query.

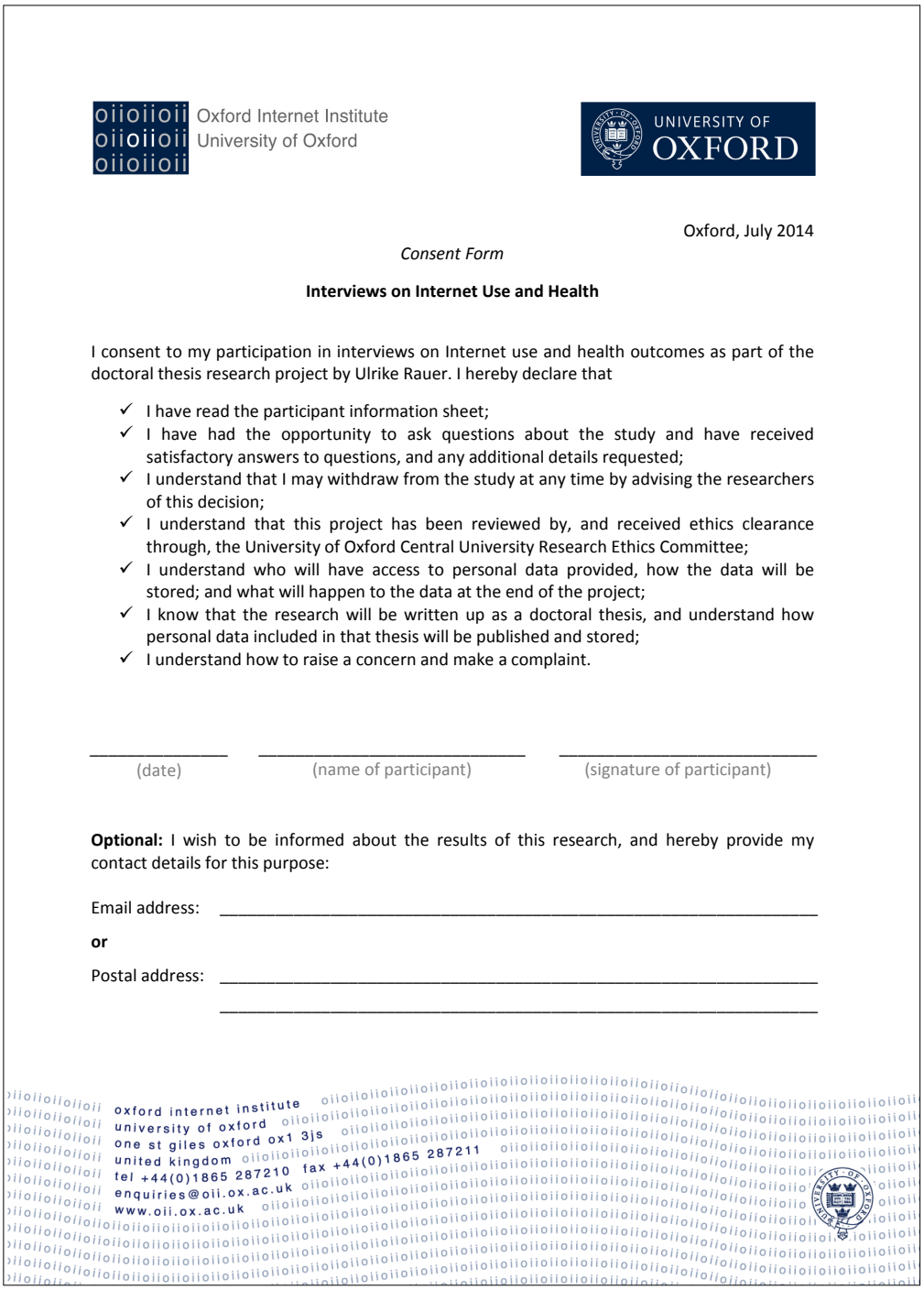
The University of Oxford Central University Research Ethics Committee has reviewed and granted ethical clearance to this study. However, if you wish to make a formal complaint, please contact the Research Ethics Committee at the University of Oxford (ethics@socsci.ox.ac.uk; 0044 1865 614871; Social Sciences & Humanities Inter-Divisional Research Ethics Committee, Oxford University, Hayes House, 75 George Street, Oxford, OX1 2BQ, UK).

Best wishes,

Ulrike Rauer

oxford internet institute
university of oxford
one st giles oxford ox1 3js
united kingdom
tel +44(0)1865 287210 fax +44(0)1865 287211
enquiries@oii.ox.ac.uk
www.oii.ox.ac.uk

Participant information. Each interviewee received this information sheet before the interview and could keep it afterwards. For the two telephone interviews, the main points were read out to the participants. Non-OxIS interviewees received a version of the participant information on which the second sentence was replaced by "For this research, you are kindly asked to take part in an interview."



Oxford Internet Institute
University of Oxford

UNIVERSITY OF OXFORD

Oxford, July 2014

Consent Form

Interviews on Internet Use and Health

I consent to my participation in interviews on Internet use and health outcomes as part of the doctoral thesis research project by Ulrike Rauer. I hereby declare that

- ✓ I have read the participant information sheet;
- ✓ I have had the opportunity to ask questions about the study and have received satisfactory answers to questions, and any additional details requested;
- ✓ I understand that I may withdraw from the study at any time by advising the researchers of this decision;
- ✓ I understand that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee;
- ✓ I understand who will have access to personal data provided, how the data will be stored; and what will happen to the data at the end of the project;
- ✓ I know that the research will be written up as a doctoral thesis, and understand how personal data included in that thesis will be published and stored;
- ✓ I understand how to raise a concern and make a complaint.

_____ (date) _____ (name of participant) _____ (signature of participant)

Optional: I wish to be informed about the results of this research, and hereby provide my contact details for this purpose:

Email address: _____

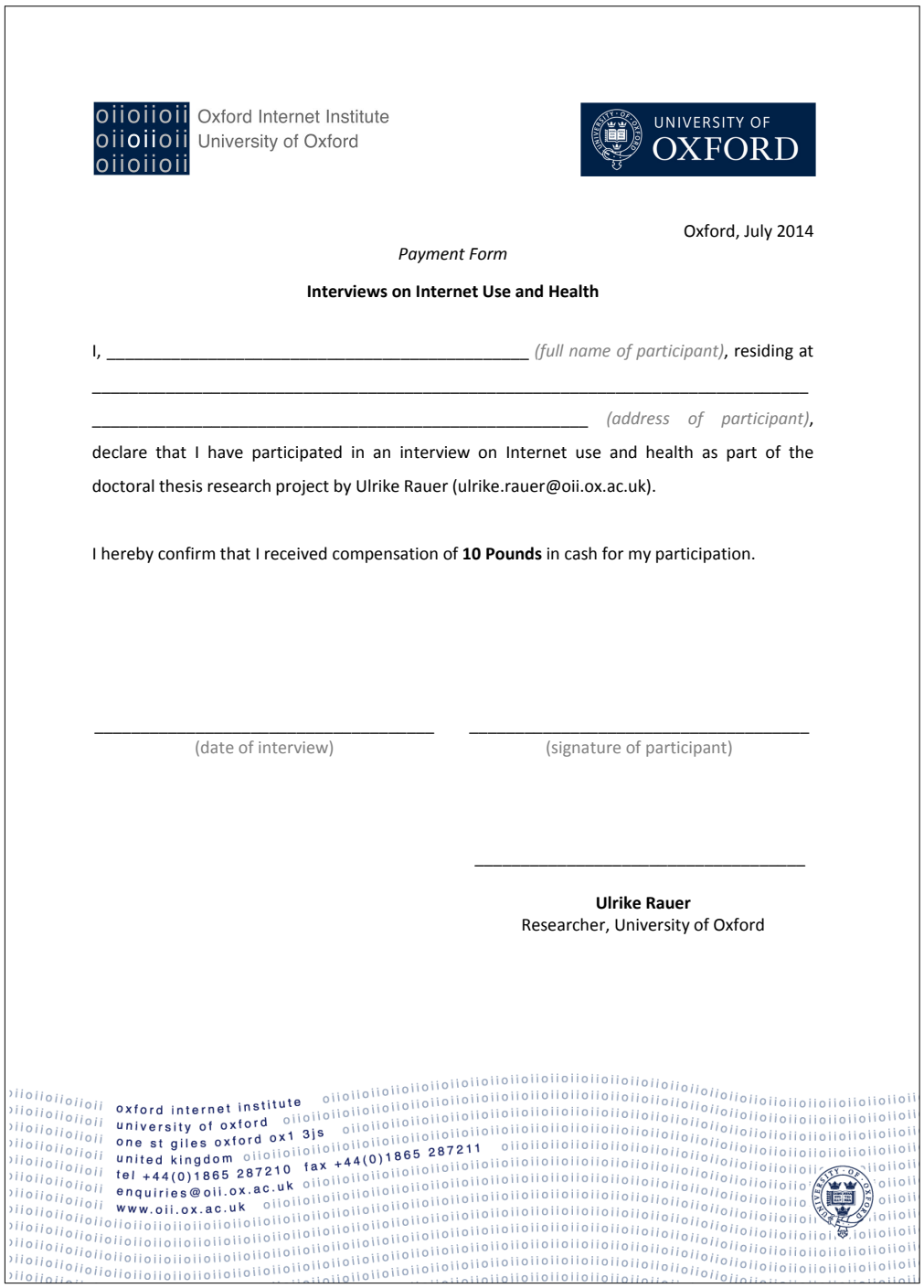
or

Postal address: _____

_____

oxford internet institute
university of oxford
one st giles oxford ox1 3js
united kingdom
tel +44(0)1865 287210 fax +44(0)1865 287211
enquiries@oii.ox.ac.uk
www.oii.ox.ac.uk

Consent form. Each interviewee signed this consent form; in the case of the two telephone interviews, consent was obtained orally. Individuals were informed about the results of the research only if they provided their email or postal address explicitly for this purpose.



Oxford Internet Institute
University of Oxford

UNIVERSITY OF OXFORD

Oxford, July 2014

Payment Form

Interviews on Internet Use and Health

I, _____ *(full name of participant)*, residing at

_____

_____ *(address of participant)*,

declare that I have participated in an interview on Internet use and health as part of the doctoral thesis research project by Ulrike Rauer (ulrike.rauer@oii.ox.ac.uk).

I hereby confirm that I received compensation of **10 Pounds** in cash for my participation.

_____ (date of interview)

_____ (signature of participant)

_____

Ulrike Rauer
Researcher, University of Oxford

oxford internet institute
university of oxford
one st giles oxford ox1 3js
united kingdom
tel +44(0)1865 287210 fax +44(0)1865 287211
enquiries@oii.ox.ac.uk
www.oii.ox.ac.uk

Payment form. If the interviewee chose the 10 GBP in cash (as opposed to donations), they filled out the payment form. Interviewees obtained the 10 GBP in an envelope with a note that read "Thank you very much for participating in the interview. I highly appreciated your insights and thoughts. Please find attached a small token of my appreciation for your time."

XVII Interview guide

Before the interview, some friendly chat not related to the research helped to relax the interviewee and build up trust (Rubin and Rubin, 2004). Recording started after the interviewee had signed the informed consent form.

Step	Interviewer questions/probing topics	Concept	Rationale
Introductory explanations and participant consent	Many thanks again for your time, I really appreciate it. I will ask you a few questions about your experiences regarding the Internet and health, and would like to hear *your* views on these – there are no right or wrong answers. This information sheet provides you with more details: Please take your time to read it and let me know if you have any questions.	–	Explain procedure and subjects to be covered (Gray, 2009); reassure that there are no "incorrect" answers (Rubin and Rubin, 2004)
	I would like to audio-tape our conversation in order to transcribe our conversation and make sure that your views are correctly represented – is that okay? Our conversation is completely confidential and anonymous – I will just use your gender, age, and *[your local authority district]* when writing up the research.	–	Ask for permission to audio-tape and guarantee confidentiality (Gray, 2009)
	If you are happy to participate, please sign this informed consent form. You only need to fill in your contact details if you are interested in hearing about my findings. It will take some time to transcribe and analyse the interviews; I will be able to them to you next summer.	–	Obtain contact details for member checking (email or postal address)
	Do you have any questions before we start?	–	Provide opportunity for questions (Gray, 2009)
Breaking the ice	I have never been to *[this area]*, please tell me a bit about the area. What's it like to live here? *[probe for health services, transport, shopping]*	Geographic environment	Obtain information on area; start with something non-personal to break the ice and build trust (Rubin and Rubin, 2004)
	Were you born here, or did you move to this place? *[probe for reasons]*	Social environment	
	How well do you know other people around here? Where is your close family?		
Understanding general Internet use and non-use	Do you use the Internet? What kind of devices do you have? *[probe for laptop, mobile, tablet]*	Online activities	Understanding user type based on different individual conceptualisations of Internet use in interviewee's own words; generate information to pick up on again throughout conversation
	– If so, since when, and what do you use it for? *[probe for communication, information]*	Reasons for non-use	
	– If not, have you used it in the past? Why are you not *[or stopped]* using the Internet? *[probe for health-related reasons]*	Motivations Enablers	
	Do you feel confident using the Internet *[or a device]*? How would you judge your skills?	Health-rel. use barriers	
	Have you asked someone to help you with using the Internet in the past? Have you ever asked someone else to look something up for you online? If so, who did you ask?	Social environment User concepts	

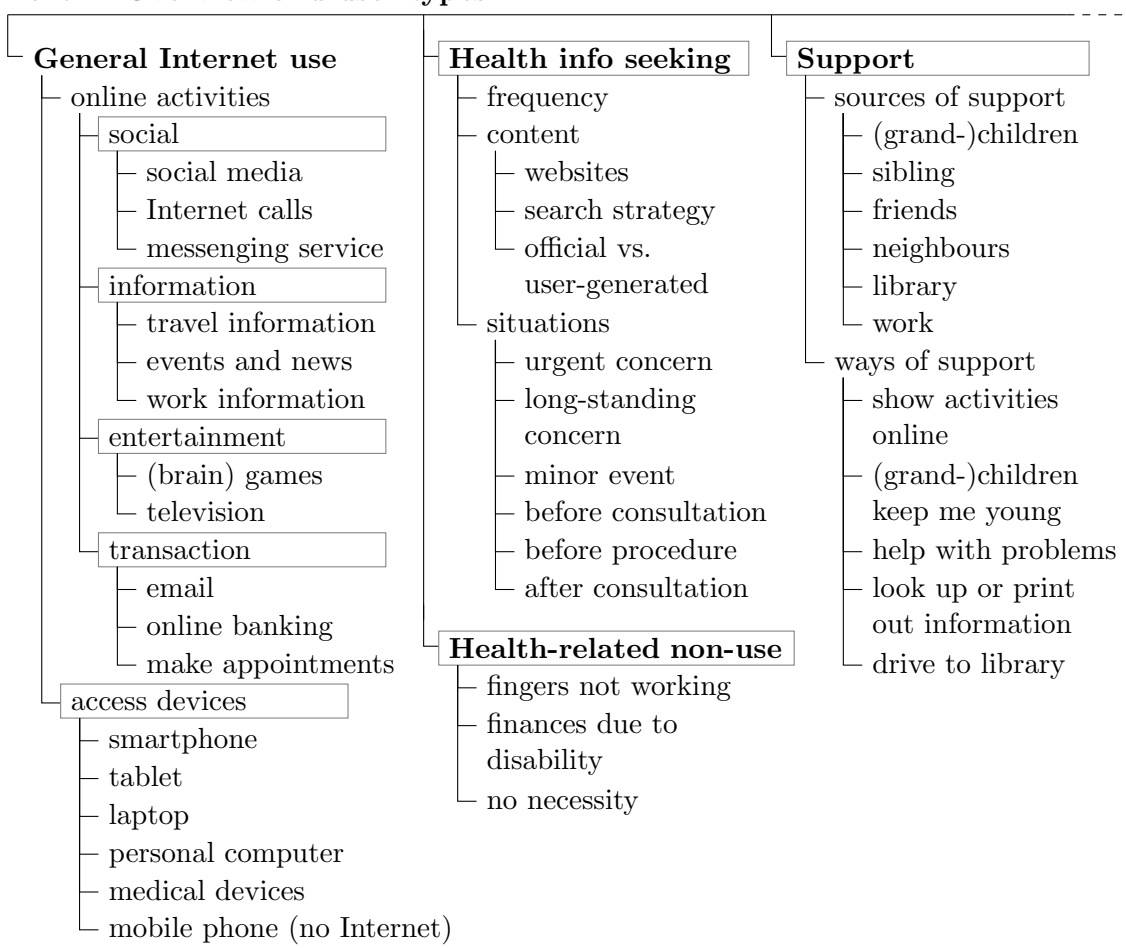
Step	Interviewer questions/probing topics	Concept	Rationale
Health-related Internet use	So have you used *[asked someone else to use]* the Internet for any health-related purposes?	Motivations Health info. seeking	Starting broadly about health-related Internet use and understand basic attitudes
	People I have spoken to have quite different views about whether the Internet is a good or bad place to search for information. What do you think?		
Specific health-related situations *Repeated several times for different situations*	Can you recall the last time when you or someone else you know had a health-related question? Please take me back to this situation and describe it as best as you can remember. *[probing for situations after/before consultation]* – In this situation, how did you find out information? *[If searched online, probe for websites, discussion forums, social media]* – Has your doctor asked about information found online or recommended a website? – Did you find the information you obtained *[online]* helpful or not?	Health info. seeking Internet outcomes	Select different specific situations that would provide vividness, richness, nuance, depth and detail about the topic (Rubin and Rubin, 2004)
	– In this situation, did the information you found online affect your decision to see a doctor about it? Did you take the information you found online to the consultation? – In this situation, how did the information you found online make you feel?	Health service use Perceived health	Understand the effects on the two outcome concepts for these specific situations
Open questions on outcome concepts	In your view, what are the most important things that affect your health? *[probe for individual/area-based factors]* What is most important for you in order to stay healthy?	Health service use Perceived health	Put Internet into perspective; understand self-efficacy, learning/ technology attitudes and trust in doctors as relevant constructs from first quantitative analysis
	How do you think is healthcare going to develop in the next years? *[probe for desired changes]*		
	Do you trust doctors? Why (not)?	Motivations	
	How interested are you in learning more about your health?		
Direct questions on outcome concepts	How would you rate your health in general on a scale from 1-5, with 5 being very good and 1 being very bad?	Perceived health	Obtain baseline values for outcome concepts; ask directly about the outcome concepts after discussing parts of it first (Rubin and Rubin, 2004)
	Do you have any long-term health conditions? *[if not mentioned before]*	Long-term health cond. Internet outcomes	
	We already talked about this, but let me ask you more generally: Does the Internet affect your health or how you perceive your health? If so, in what ways?		
	How many times in the last year have you seen your GP? How many times have you been to hospital during the last year?	Health service use Internet outcomes	
	We have also talked about this already, but let me ask you directly: does the Internet affect how often you see a GP or go to hospital? If so, in what ways?		

Step	Interviewer questions/probing topics	Concept	Rationale
Other interviewee reflections and debrief	Is there anything else I haven't asked about that you think may be relevant?	–	Obtain other thoughts on interviewees mind (Arksey and Knight, 1999)
	Thank you very much, you have really helped me a lot to better understand the relationship between Internet and health. This is also what my research is about *[providing more information on research questions]*	–	
Socio-demographic information	I would just like to ask you a few last questions about yourself. *[only ask the ones that the interviewee did not already bring up himself/herself]* – Please tell me a bit about what you do for a living *[probe for occupation, role, size of organisation]* – What's the highest educational qualification you have? – When were you born? *[only if seems to deviate from the one in OxIS, or for non-OxIS participants]*	Age Gender NS-SeC Education	Obtain socio-demographic information of relevance to the research; ask potentially sensitive questions relatively late in order to minimise dropout (Arksey and Knight, 1999)
Next steps	Many thanks again for your participation; it was really valuable to hear your views. Our conversation will be transcribed, everything we discussed is anonymous and confidential. I will then analyse all the interviews and write up the findings.	–	Thank the interviewee (again), mention value of results to leave interviewee with positive feeling, explain next steps and follow-up work (Gray, 2009).
	If I have any further questions, I would like to briefly contact you if that is ok, and feel free to contact me at any time if you have any further questions – my contact details are on the participant information sheet. *[If applicable]* As discussed, I will send you an overview of the results of the research next summer.	–	
	As promised, you get 10 Pounds as compensation for your time, or you can donate it to a charity of your choice.	–	Allow interviewee to pick cause close to their heart; provide further topics for discussion based on chosen charity

XVIII Coding scheme

Below is the coding scheme that has been developed in several coding cycles, with new codes being added, as well as the "lumping and splitting" of codes (Saldaña, 2012). Predetermined codes help to ensure comparability across the qualitative and the quantitative part and answer the research questions. These are marked by the codes in boxes (□). All lower level codes emerged from the qualitative interview data.

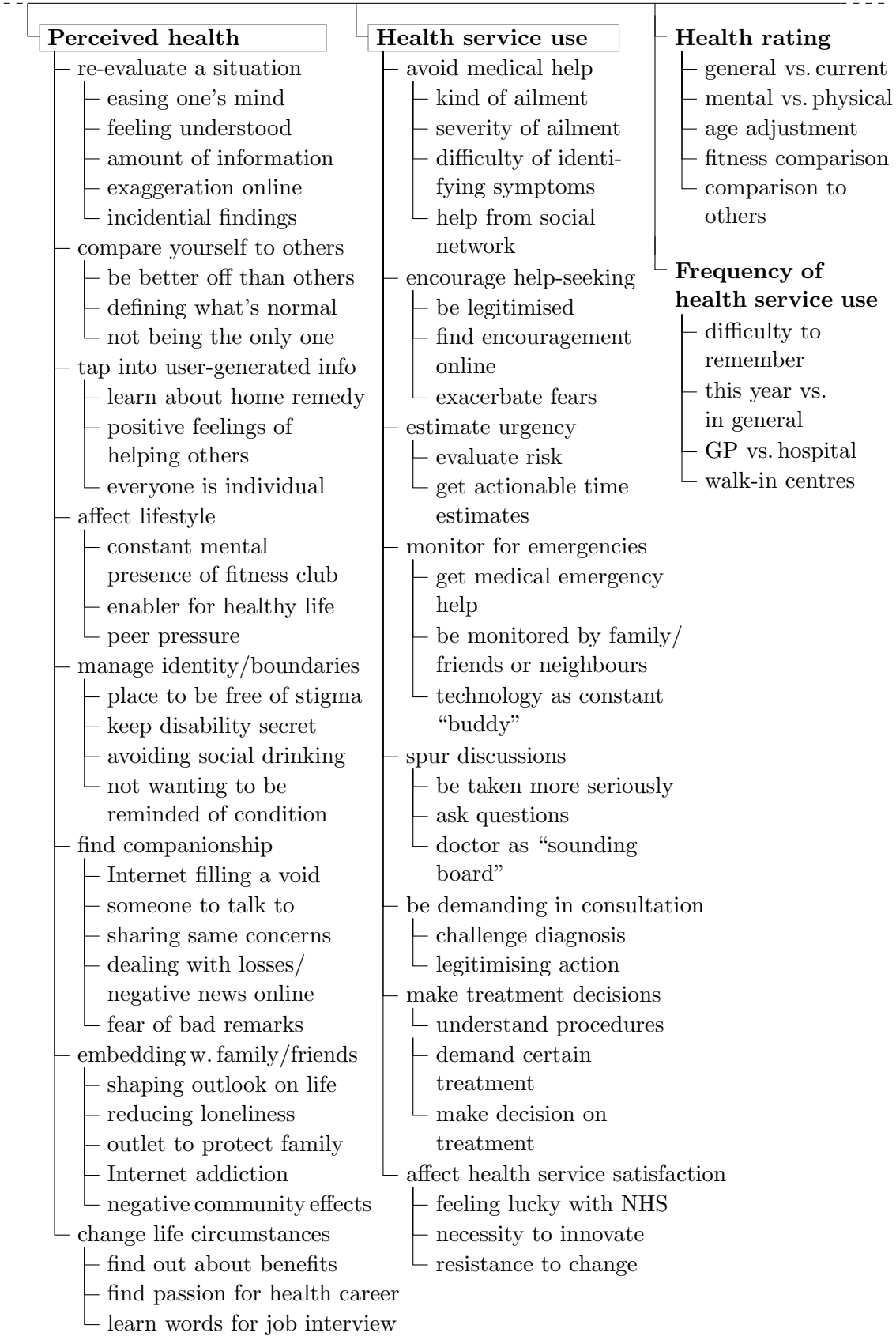
Part 1: Overview and user types



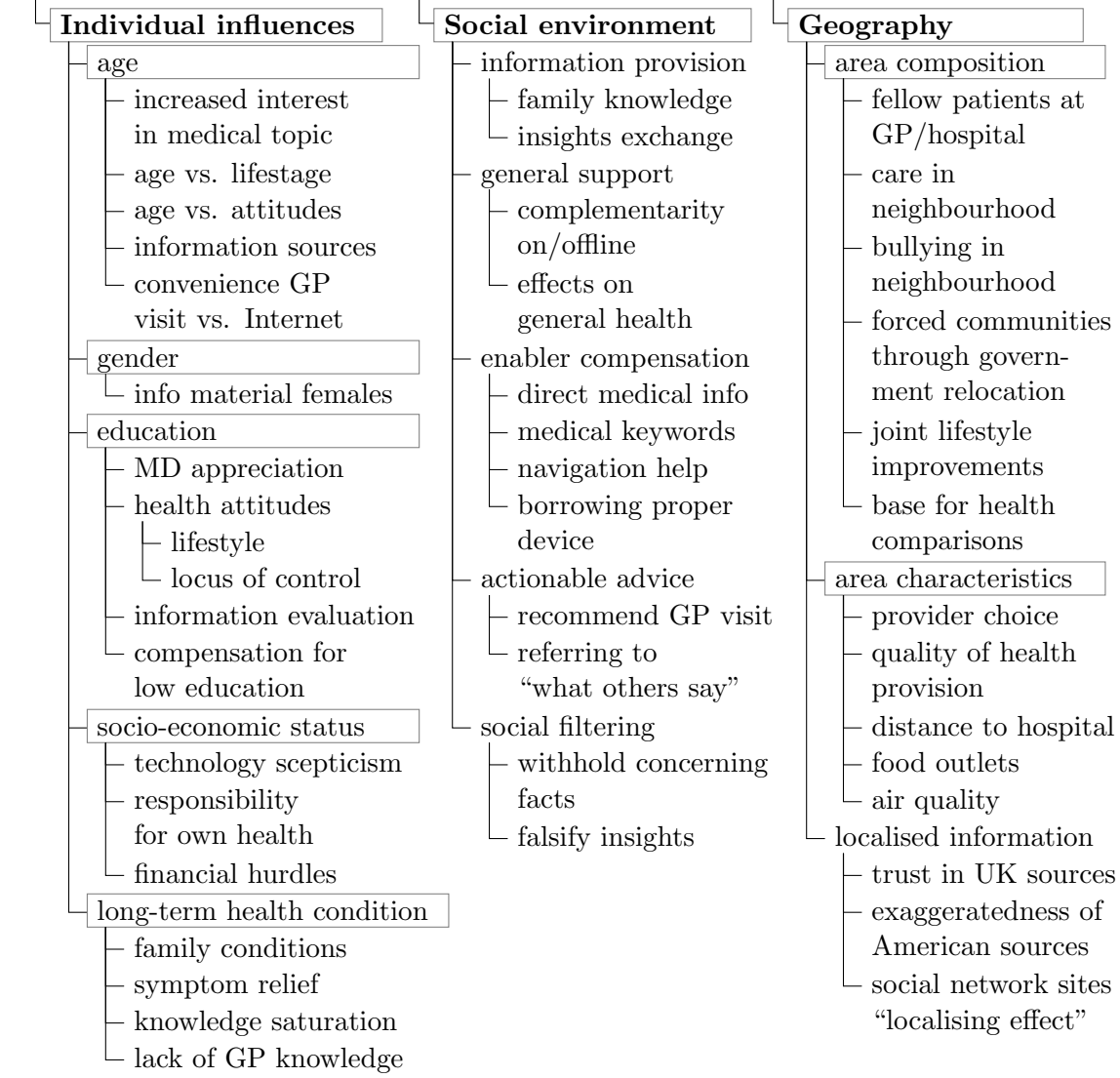
Part 2: Antecedent factors



Part 3: Influences on perceived health and health service use



Part 4: Individual and contextual influences

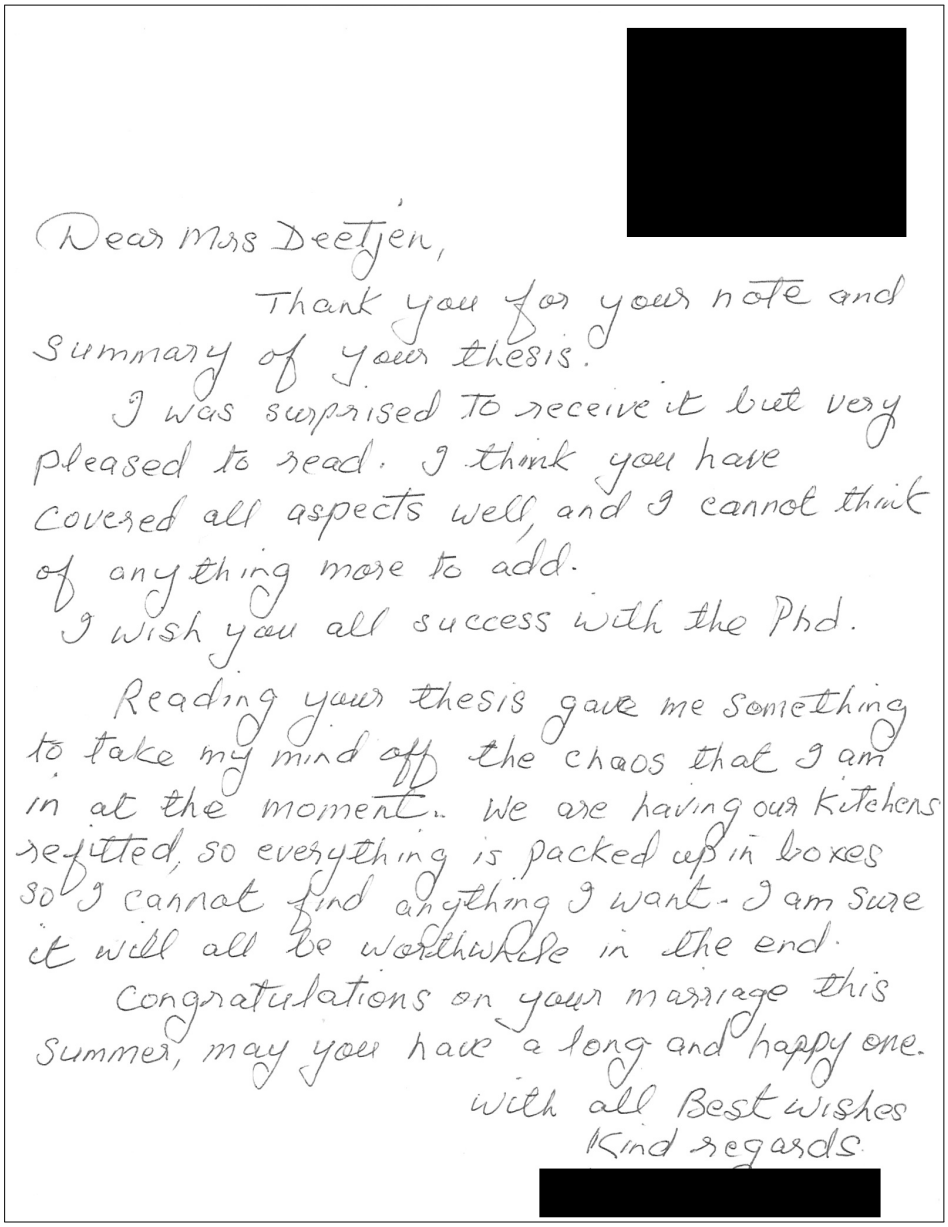


Part 5: Other codes

- **Health conditions**
  - Conditions
    - ADHD
    - alcoholism
    - arthritis
    - Asperger's syndrome
    - asthma
    - autism
    - back pains
    - cataract
    - COPD
    - dementia
    - depression
    - diabetes
    - dieting/weight
    - disintegrated spine/fracture
    - DVT
    - eczema
    - epilepsy
    - gastric cancer
    - glaucoma
    - hearing loss
    - heart failure
    - hypertension
    - hypothyroidism
    - IBS
    - men's health
    - MS
    - OCD
    - pancreatic cancer
    - Parkinson's disease
    - pregnancy
    - rare chromosome disorder
    - septal defects
    - shortsightedness
    - skin cancer
    - smoking
    - stroke
- **Recommendations**
  - health professionals
    - recommend specific websites
    - use simpler language
    - give actionable advice
  - policymakers
    - mark official pictures
    - provide clear domain names
    - have commercials on television on online health information
- **Methodology**
  - interview location
    - protected home
    - home due to health reasons
  - interview mode
    - face-to-face vs. telephone
  - thought-provoking prompts
- **Expressions**
  - Internet
    - Internet vs. Facebook
    - Internet vs. NHS website
    - Internet as Safari
  - health
    - GP as quack

XIX Research summary and interviewee responses

Kathleen:


[redacted]

Dear Miss Deetjen,

Thank you for your note and summary of your thesis.

I was surprised to receive it but very pleased to read. I think you have covered all aspects well, and I cannot think of anything more to add.

I wish you all success with the Phd.

Reading your thesis gave me something to take my mind off the chaos that I am in at the moment.. We are having our kitchens refitted, so everything is packed up in boxes so I cannot find anything I want - I am sure it will all be worthwhile in the end.

Congratulations on your marriage this summer, may you have a long and happy one.

With all Best Wishes
Kind regards

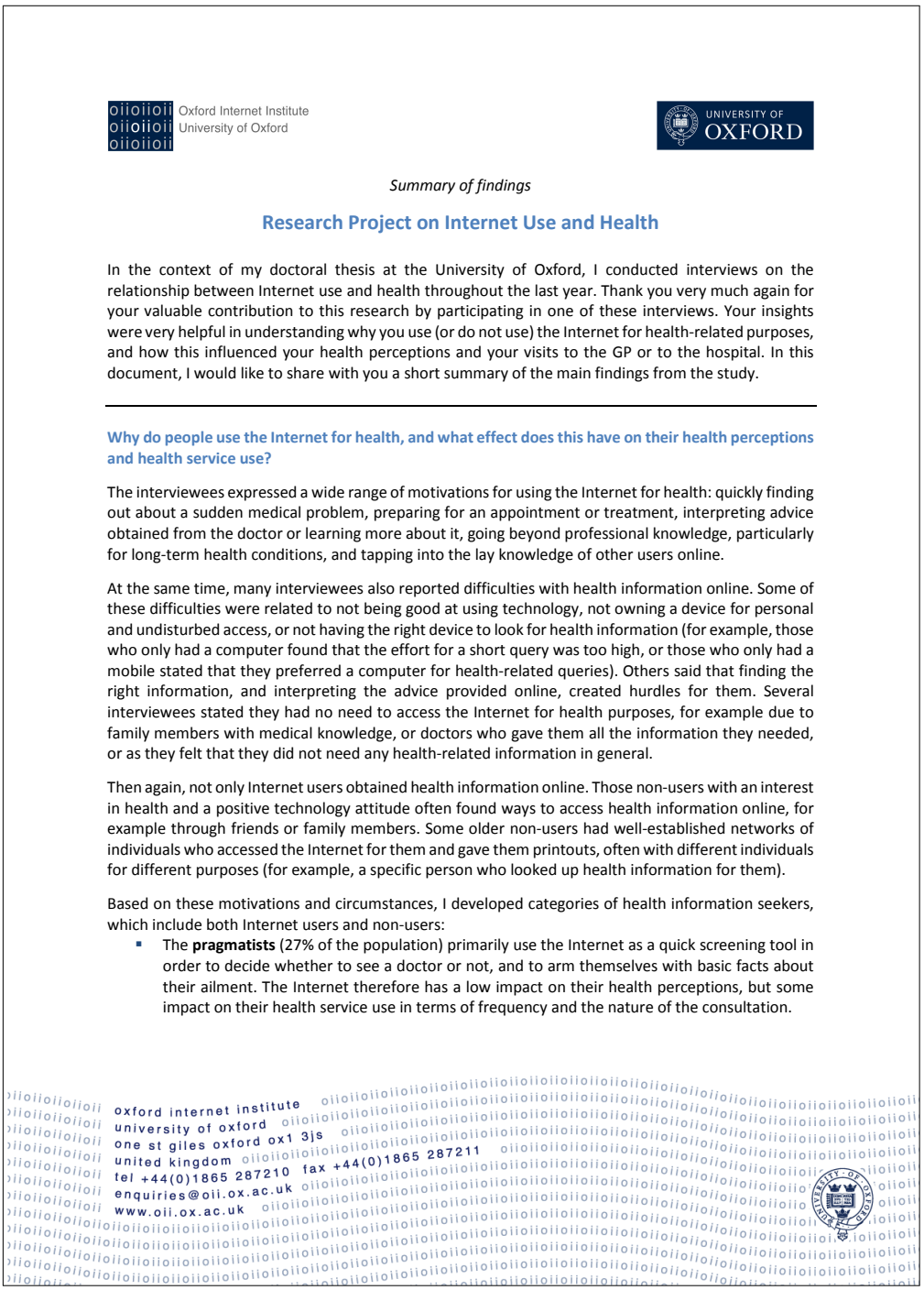
[redacted]

Ian, mentioning his wife **Isabelle**:

Thankyou for your very kind words. I'm very pleased that i could contribute information to help aid you in your career. [redacted] was also very pleased. I wish you every success that you want to achieve in life. good luck. [redacted]

Research Summary

These three pages were sent to every interviewee who had provided his/her email address or postal address on the consent form, in addition to a personal hand-written note or a personalised email that highlighted a unique contribution of the interviewee or referred to some specific circumstances about the interview.



oiioiioii Oxford Internet Institute
oiioiioii University of Oxford
oiioiioii

UNIVERSITY OF OXFORD

Summary of findings

Research Project on Internet Use and Health

In the context of my doctoral thesis at the University of Oxford, I conducted interviews on the relationship between Internet use and health throughout the last year. Thank you very much again for your valuable contribution to this research by participating in one of these interviews. Your insights were very helpful in understanding why you use (or do not use) the Internet for health-related purposes, and how this influenced your health perceptions and your visits to the GP or to the hospital. In this document, I would like to share with you a short summary of the main findings from the study.

Why do people use the Internet for health, and what effect does this have on their health perceptions and health service use?

The interviewees expressed a wide range of motivations for using the Internet for health: quickly finding out about a sudden medical problem, preparing for an appointment or treatment, interpreting advice obtained from the doctor or learning more about it, going beyond professional knowledge, particularly for long-term health conditions, and tapping into the lay knowledge of other users online.

At the same time, many interviewees also reported difficulties with health information online. Some of these difficulties were related to not being good at using technology, not owning a device for personal and undisturbed access, or not having the right device to look for health information (for example, those who only had a computer found that the effort for a short query was too high, or those who only had a mobile stated that they preferred a computer for health-related queries). Others said that finding the right information, and interpreting the advice provided online, created hurdles for them. Several interviewees stated they had no need to access the Internet for health purposes, for example due to family members with medical knowledge, or doctors who gave them all the information they needed, or as they felt that they did not need any health-related information in general.

Then again, not only Internet users obtained health information online. Those non-users with an interest in health and a positive technology attitude often found ways to access health information online, for example through friends or family members. Some older non-users had well-established networks of individuals who accessed the Internet for them and gave them printouts, often with different individuals for different purposes (for example, a specific person who looked up health information for them).

Based on these motivations and circumstances, I developed categories of health information seekers, which include both Internet users and non-users:

- The **pragmatists** (27% of the population) primarily use the Internet as a quick screening tool in order to decide whether to see a doctor or not, and to arm themselves with basic facts about their ailment. The Internet therefore has a low impact on their health perceptions, but some impact on their health service use in terms of frequency and the nature of the consultation.

oxford internet institute
university of oxford
one st giles oxford ox1 3js
united kingdom
tel +44(0)1865 287210 fax +44(0)1865 287211
enquiries@oii.ox.ac.uk
www.oii.ox.ac.uk

- The **learners** (22%) have a particular interest in finding out everything about their and others' health before and after consultations, and frequently establish themselves as lay experts online or in their offline social circles. For this group, the Internet makes a substantial contribution to how they perceive their health, and shapes their health service use as well.
- In contrast, for the **worriers** (10%), as the group with the highest frequency of health information seeking, the Internet leads to more concerns and, in some instances, higher health service use. As a consequence, some of the worriers have stopped looking for health information online, or only look at official information (for example, on the NHS website).
- The **sceptics** (14%), with a lower trust in doctors, embrace the Internet as a way of avoiding health service use, and see the Internet as a way to ensure that the doctor is reaching the right judgement. Thus, the Internet has an impact on how often they see a doctor, and makes them more demanding in health service use.
- The **delegators** (9%) value the convenience associated with other people finding out health information for them; whether they are Internet users or non-users does not matter in this regard. While these social filtering mechanisms lead to lower concerns based on health information online, there is nearly no influence on the frequency of health service use.
- Finally, the **adigitals** (18%) often do not use the Internet, and if they do, many have strong views that the Internet should not be used for health services and health information at all. As a result, adigitals do not perceive any effects on their health and how they make use of health services.

How does the Internet affect people's health perception and health service use?

A closer look at the effects of Internet use revealed a range of different ways in which the Internet affected how individuals perceived their health, and how they made use of health services.

For **health service use**, several interviewees said they managed to get by without medical help because of the information found online. In other cases, while they still contacted a medical professional, the Internet helped them to estimate the urgency of their condition. At the same time, the Internet also helped others to feel normal about having a certain condition, and feel justified in seeking medical help, particularly for conditions that they felt embarrassed about. Several interviewees said that both the telephone and the Internet helped their families and friends to check on them, and contact medical help if necessary. In addition, several interviewees found that the Internet did not necessarily change how often they sought medical help, but rather shaped how the consultation went: For example, information found online spurred discussions in consultations, helped them to be more demanding in some cases, and allowed them to make decisions about treatments. Finally, several interviewees also said that by using the Internet they obtained a different view of the health system, for example based on discussions with individuals from other countries.

The Internet also affected individuals' **perceived health** as they re-evaluated a situation based on information found online or used user-generated health tips. Many interviewees highlighted that the Internet helped them to compare themselves to others – either those who were feeling better after a certain treatment, giving them confidence for the way forward, or those who were much worse off, helping them to put their own complaints into perspective. Some interviewees explained that the Internet affected their perceptions as it provided more information about lifestyles during everyday life. But others thought of the Internet as a barrier between different parts of their life (for example, as a

oxford internet institute
university of oxford
one st giles oxford ox1 3js
united kingdom
tel +44(0)1865 287210 fax +44(0)1865 287211
enquiries@oii.ox.ac.uk
www.oii.ox.ac.uk

space where one can be free of the disease, as a way to meet other people without being forced into social drinking, or as a way to prevent a professional network from knowing about a certain health condition). Finally, in some instances, the Internet also provided companionship, helped individuals to feel closer to family and friends, and was seen as a way to change life circumstances more generally (for example, through learning words for job interviews or learning about available benefits after an accident). All of these were seen as potential indirect influences on perceived health.

A general observation was that those for whom the Internet was highly embedded in their everyday lives tended to rate their health as slightly lower. They explained this lower rating based on not getting as much exercise as they would like to, or not eating healthily enough. This indicated that the Internet shifted the yardstick of what "perfect health" meant, as the Internet creates a higher focus on lifestyles and healthy living than the offline realm.

Finally, of particular importance was the individuals' social environment. While some people were able to compensate for a lack of skills in accessing the Internet with the help of family members and friends, the social environment also provided information, helped to generate advice they could act upon, and provided general support that was seen as beneficial in the context of health. Family members with medical knowledge (and doctors in general) sometimes served as a sounding board for information from online sources. At the same time, family members looking up health information also acted as an important filter for "bad news" found online, just as the online realm provided an outlet to discuss or read up on their questions instead of worrying other family members.

I hope this short summary reflects the most important parts of our conversation last year, even though it cannot convey all the rich and detailed insights you provided me. **If you have any additional thoughts on this summary or think I did not fully capture an important part of our conversation, I would be grateful to hear back from you and to include your comments in the results of this research.** I can be reached via email (ulrike.deetjen@oii.ox.ac.uk) or in writing (Ulrike Deetjen, Oxford Internet Institute, 1 St Giles, Oxford OX1 3JS) and would be very happy to hear from you.

I am currently in the process of carrying out further analysis of quantitative data from the Oxford Internet Surveys (I obtained your contact details from these surveys), as well as census data and hospital episode statistics. Moreover, I have nearly finished writing up the results as a doctoral thesis including more details and quotes from our interview. This thesis will be deposited with the University of Oxford so that it can be used in future research. If you would like a copy of the document please let me know.

The findings will also be published and included in presentations. These will be targeted at NHS programmes such as the "Widening Digital Participation" project, and other government initiatives around digital inclusion and digital health. **As discussed during the interview, you will never be personally identified in any of the resulting documents.**

Thank you very much again for taking part in the interview. I really appreciated your time and the valuable insights for my research project. Please do not hesitate to contact me with any questions, concerns or comments you may have.

All the best,

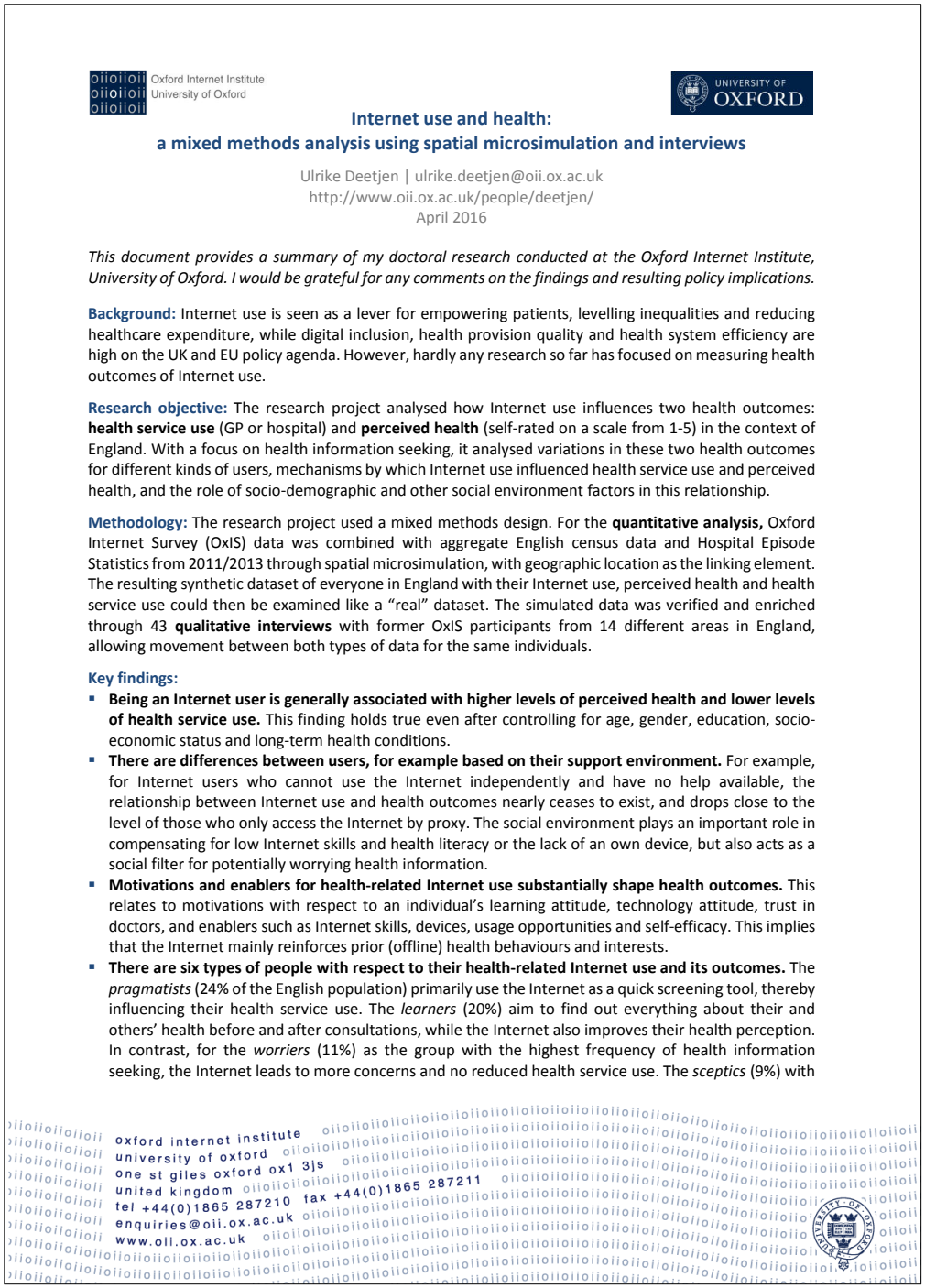
Ulrike Deetjen

oxford internet institute
university of oxford
one st giles oxford ox1 3js
united kingdom
tel +44(0)1865 287210 fax +44(0)1865 287211
enquiries@oii.ox.ac.uk
www.oii.ox.ac.uk

Note: The addition "(I obtained your contact details from these surveys)" in the fifth paragraph was taken out in the version sent to non-OxIS participants.

XX Policy summary

This two-page summary was published on the OII website, and sent to decision-makers at NHS England, Public Health England and the European Commission.



Oxford Internet Institute
University of Oxford

UNIVERSITY OF OXFORD

Internet use and health:
a mixed methods analysis using spatial microsimulation and interviews

Ulrike Deetjen | ulrike.deetjen@oii.ox.ac.uk
http://www.oii.ox.ac.uk/people/deetjen/
April 2016

This document provides a summary of my doctoral research conducted at the Oxford Internet Institute, University of Oxford. I would be grateful for any comments on the findings and resulting policy implications.

Background: Internet use is seen as a lever for empowering patients, levelling inequalities and reducing healthcare expenditure, while digital inclusion, health provision quality and health system efficiency are high on the UK and EU policy agenda. However, hardly any research so far has focused on measuring health outcomes of Internet use.

Research objective: The research project analysed how Internet use influences two health outcomes: **health service use** (GP or hospital) and **perceived health** (self-rated on a scale from 1-5) in the context of England. With a focus on health information seeking, it analysed variations in these two health outcomes for different kinds of users, mechanisms by which Internet use influenced health service use and perceived health, and the role of socio-demographic and other social environment factors in this relationship.

Methodology: The research project used a mixed methods design. For the **quantitative analysis,** Oxford Internet Survey (OxIS) data was combined with aggregate English census data and Hospital Episode Statistics from 2011/2013 through spatial microsimulation, with geographic location as the linking element. The resulting synthetic dataset of everyone in England with their Internet use, perceived health and health service use could then be examined like a "real" dataset. The simulated data was verified and enriched through 43 **qualitative interviews** with former OxIS participants from 14 different areas in England, allowing movement between both types of data for the same individuals.

Key findings:

- **Being an Internet user is generally associated with higher levels of perceived health and lower levels of health service use.** This finding holds true even after controlling for age, gender, education, socio-economic status and long-term health conditions.
- **There are differences between users, for example based on their support environment.** For example, for Internet users who cannot use the Internet independently and have no help available, the relationship between Internet use and health outcomes nearly ceases to exist, and drops close to the level of those who only access the Internet by proxy. The social environment plays an important role in compensating for low Internet skills and health literacy or the lack of an own device, but also acts as a social filter for potentially worrying health information.
- **Motivations and enablers for health-related Internet use substantially shape health outcomes.** This relates to motivations with respect to an individual's learning attitude, technology attitude, trust in doctors, and enablers such as Internet skills, devices, usage opportunities and self-efficacy. This implies that the Internet mainly reinforces prior (offline) health behaviours and interests.
- **There are six types of people with respect to their health-related Internet use and its outcomes.** The *pragmatists* (24% of the English population) primarily use the Internet as a quick screening tool, thereby influencing their health service use. The *learners* (20%) aim to find out everything about their and others' health before and after consultations, while the Internet also improves their health perception. In contrast, for the *worriers* (11%) as the group with the highest frequency of health information seeking, the Internet leads to more concerns and no reduced health service use. The *sceptics* (9%) with

oxford internet institute
university of oxford
one st giles oxford ox1 3js
united kingdom
tel +44(0)1865 287210 fax +44(0)1865 287211
enquiries@oii.ox.ac.uk
www.oii.ox.ac.uk

a lower trust in doctors embrace the Internet as a way of verifying information from medical professionals or avoiding health service use. The *delegators* (11%) prefer to have others find out health information for them, and while these social filtering mechanisms avoid concerns based on health information online, there is nearly no influence on the frequency of health service use. Finally, for the *adigitals* (25%), there are no effects on either outcome.

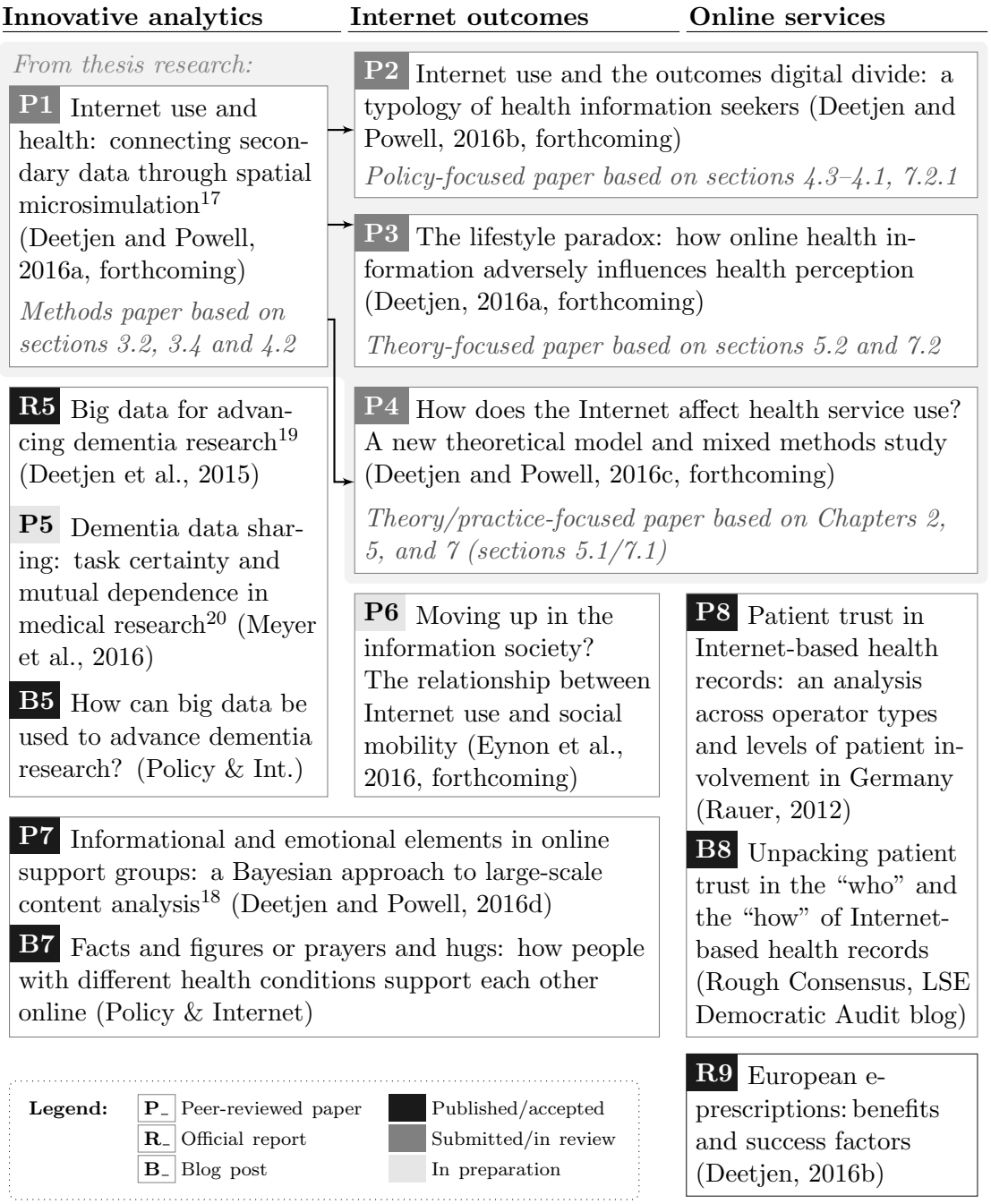
- **The Internet affects health service use and perceived health by various mechanisms.** The Internet reduces *health service use* or encourages it for stigmatised conditions, affects individuals' sense of urgency, shapes the discussion in the consultation and satisfaction with the NHS in general. For *perceived health*, the Internet helps to re-evaluate certain situations, makes individuals feel normal about their condition, creates a space where they can briefly forget about their disease, provides companionship, and changes life circumstances more generally, with positive effects on health perception. However, this research also provided evidence for the "lifestyle paradox": based on reading about health online, younger individuals in particular shift their yardsticks for what perfect health means in terms of fitness, thereby lowering health perception despite being in good health.
- **The data supports a shift from usage to outcome inequalities.** While it has become widely accepted that digital inequalities have shifted from access to usage with growing Internet penetration, the next level of inequalities relates to the benefits that individuals reap from using the technology.

Implications for policy and practice:

- **Existing NHS initiatives are important for creating Internet-based health benefits and supporting efficiency in the health system while leaving nobody behind.** Programmes such as "Widening Digital Participation" address an important skills dimension that was evident in this research, and particularly help those individuals who already attribute some usefulness to the Internet, but cannot use the Internet due to a lack of enablers. Moreover, some ongoing activities – such as providing certificates of trustworthiness for apps to make them "NHS-approved" – will be highly welcomed by those who form part of the worrier and the pragmatist types.
- **Enabling Internet-based health outcomes and introducing new digital services requires understanding different types of individuals.** The pragmatists, learners and some of the sceptics valued the efficiency associated with digital services, whereas this was not the case for the adigitals. Understanding these differences, tailoring interventions that account for the different motivations, and addressing all enablers including interest and usage opportunities – rather than focusing on skills and devices only – is paramount.
- **Practitioners need to actively embrace the benefits that the Internet may entail.** Many patients wish to talk openly about their activities online, and use the doctor as a trusted advisor and "sounding board" for information found online. Moreover, the worriers and the delegators in particular would appreciate specific recommendations for online health resources for realising Internet-based health outcomes. Adigitals equate the shift towards digital services with a deterioration of healthcare, so it is crucial that practitioners are champions in bringing the technology to patients.
- **Social inclusion remains an important policy priority for Internet-based health outcomes.** While social inclusion itself has positive effects on health, this research showed its importance for offsetting skills gaps, and as a social filtering mechanism for obtaining health information. At the same time, the Internet itself may help users to feel more integrated with family and friends and enable the finding of companionship – with positive effects on their health perception.
- **Aggregate health data should be made more accessible so that value can be derived from it.** While data privacy is important and health data should be protected to avoid endangering individuals' trust in large-scale data resources, data provision guidelines should be broadened to reflect the shifting boundaries of medically relevant data beyond the confines of the health system, and embrace innovative approaches such as spatial microsimulation (used in this research) as a cost-effective and privacy-preserving way of creating novel insights from existing data.

oxford internet institute
university of oxford
one st giles oxford ox1 3js
united kingdom
tel +44(0)1865 287210 fax +44(0)1865 287211
enquiries@oii.ox.ac.uk
www.oii.ox.ac.uk

XXI Overview of publications



[17] Presented at the 6th ESRC Research Methods Festival, Oxford, July 2014.

[18] Presented at the BSA Medical Sociology Conference in Birmingham, September 2014.

[19] Presented at OECD conference in Toronto, Ontario in September 2014, and the WHO First Ministerial Conference on Global Action Against Dementia, Geneva, March 2015.

[20] Presented at the Society for Social Studies of Science, Denver, Colorado, November 2015.

Bibliography

Adelman, M. B., M. R. Parks, and T. L. Albrecht (1987). Beyond close relationships: support in weak ties. In T. L. Albrecht and M. B. Adelman (Eds.), *Communicating Social Support*, pp. 126–147. Newbury Park, CA: SAGE.

Adler, N. E. and K. Newman (2002). Socioeconomic disparities in health: pathways and policies. *Health Affairs 21*(2), 60–76.

Adler, P. S. and S.-W. Kwon (2000). Social capital: the good, the bad and the ugly. In E. L. Lesser (Ed.), *Knowledge and Social Capital: Foundations and Applications*, pp. 89–119. Woburn, MA: Butterworth-Heinemann.

Albrecht, G. L. and P. J. Devlieger (1999). The disability paradox: high quality of life against all odds. *Social Science & Medicine 48*(8), 977–988.

Alemi, F., M. Mosavel, R. C. Stephens, A. Ghadiri, J. Krishnaswamy, and H. Thakkar (1996). Electronic self-help and support groups. *Medical Care 34*(10), OS32–OS44.

Allen, M. (2010). The experience of connectivity. *Information, Communication & Society 13*(3), 350–374.

Alloway, R. and P. Bebbington (1987). The buffer theory of social support–a review of the literature. *Psychological Medicine 17*(01), 91–108.

Andersen, R. (1968). A behavioral model of families' use of health services. Research series 25, Chicago: Center for Health Administration Studies.

Andersen, R. M. (1995). Revisiting the behavioral model and access to medical care: does it matter? *Journal of Health & Social Behavior 36*(1), 1–10.

Andersen, R. M. and P. Davidson (2007). Improving access to care in America: individual and contextual characters. In R. M. Andersen, T. Rice, and J. Kominski (Eds.), *Changing the US Health Care System: Key Issues in Health Services Policy and Management*. San Francisco, CA: Jossey-Bass.

Anderson, B. (2007a). DCLG SimInc D1: Preliminary Results Report. DCLG SimInc Project Deliverable. Ipswich, Chimera, University of Essex.

Anderson, B. (2007b). Creating small area income estimates for England: spatial microsimulation modelling. Chimera working paper number 2007-07, University of Essex, Colchester.

Anderson, B. (2009). Creating small area income deprivation estimates for Northern Ireland: spatial microsimulation modelling. Working paper 2009-07, Centre for Research in Economic Sociology and Innovation (CRESI), University of Essex, Colchester, UK.

Anderson, B. (2013). Estimating small-area income deprivation: an iterative proportional fitting approach. In R. Tanton and K. L. Edwards (Eds.), *Spatial Microsimulation: A Reference Guide for Users*, pp. 69–85. Dordrecht: Springer.

Anderson, B. and E. Speed (2009). An evidence-based approach to digital inclusion for health. Report to Solihull Care Trust, Centre for Research in Economic Sociology and Innovation (CRESI), University of Essex, Colchester, UK.

Anderson, B. and E. Speed (2010). Social media and health: implications for primary health care providers. Report to Solihull Care Trust, Colchester, University of Essex.

Anderson, B. and K. Tracey (2008). Digital living: the impact (or otherwise) of the Internet on everyday British life. In B. Wellman and C. Haythornthwaite (Eds.), *The Internet in Everyday Life*, Information Age Series. Malden, MA: Blackwell Publishing.

Anderson, D. R., W. A. Link, D. H. Johnson, and K. P. Burnham (2001). Suggestions for presenting the results of data analyses. *The Journal of Wildlife Management 65*(3), 373–378.

Anderson, J. G., M. R. Rainey, and G. Eysenbach (2003). The impact of CyberHealthcare on the physician–patient relationship. *Journal of Medical Systems 27*(1), 67–84.

Antonovsky, A. (1979). *Health, stress, and coping: new perspectives on mental and physical well-being*. San Francisco, CA: Jossey-Bass.

Arksey, H. and P. Knight (1999). *Interviewing for Social Scientists* (1 ed.). London: SAGE Publications Ltd.

Atkinson, N. L., S. L. Saperstein, and J. Pleis (2009). Using the Internet for health-related activities: findings from a national probability sample. *Journal of Medical Internet Research 11*(1), e4.

Ayers, S. L. and J. J. Kronenfeld (2007). Chronic illness and health-seeking information on the Internet. *Health: 11*(3), 327–347.

Ayres, L., K. Kavanaugh, and K. A. Knafl (2003). Within-case and across-case approaches to qualitative data analysis. *Qualitative Health Research 13*(6), 871–883.

Babbie, E. R. (2009). *The Practice of Social Research* (12th ed.). International Edition. Belmont, CA: Wadsworth Cengage Learning.

Bakardjieva, M. (2005). *Internet society: The Internet in everyday life*. Sage.

Baker, P. (2011). Times may change, but we will always have money: diachronic variation in recent British English. *Journal of English Linguistics 39*(1), 65–88.

Ballas, D. and B. Anderson (2013). Iterative proportional fitting (IPF). In A. Whitworth (Ed.), *Evaluations and Improvements in Small Area Estimation Methodologies*, Methodological Review paper. Accessed 07 February 2014, http://eprints.ncrm.ac.uk/3210/1/sme_whitworth.pdf: National Centre for Research Methods (NCRM).

Ballas, D., G. Clarke, D. Dorling, and H. Eyre (2005). SimBritain: a spatial microsimulation approach to population dynamics. *Population, Space & Place 11*(1), 13–34.

Ballas, D., G. Clarke, D. Dorling, J. Rigby, and B. Wheeler (2006). Using geographical information systems and spatial microsimulation for the analysis of health inequalities. *Health Informatics Journal 12*(1), 65–79.

Ballas, D., D. Rossiter, B. Thomas, G. Clarke, and D. Dorling (2005). *Geography matters: simulating the local impacts of national social policies*. Joseph Rowntree Foundation.

Baltagi, B. (2008). *Econometric analysis of panel data*. John Wiley & Sons.

Bandura, A. (1986). *Social foundations of thought and action: a social cognitive theory*. Prentice-Hall series in social learning theory. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education & Behavior 31*(2), 143–164.

Bansil, P., N. L. Keenan, A. I. Zlot, and J. C. Gilliland (2006). Health-related information on the web: Results from the HealthStyles survey, 2002–2003. *Preventing Chronic Disease 3*(2), 1–10.

Barbour, R. S. (1999). The case for combining qualitative and quantitative approaches in health services research. *Journal of Health Services Research & Policy 4*(1), 39–43.

Barrett, E. (2014). Introduction. In E. Barrett and B. Bolt (Eds.), *Practice as Research: Approaches to Creative Arts Enquiry*. London; New York, NY: I. B. Tauris.

Bell, A. V. (2014). "I think about Oprah": Social class differences in sources of health information. *Qualitative Health Research 24*(4), 506–516.

Belsley, D. A., E. Kuh, and R. E. Welsh (1980). *Regression Diagnostics: identifying Influential Data and Sources of Collinearity*. New York, NY: John Wiley & Sons.

Bentley, J. and P. Thacker (2004). The influence of risk and monetary payment on the research participation decision making process. *Journal of Medical Ethics 30*(3), 293–298.

Berg, B. L. and H. Lune (2014). *Qualitative Research Methods for the Social Sciences* (8 ed.). Harlow: Pearson Education Limited.

Berger, M., T. H. Wagner, and L. C. Baker (2005). Internet use and stigmatized illness. *Social Science & Medicine 61*(8), 1821–1827.

Bergmo, T. S., P. E. Kummervold, D. Gammon, and L. B. Dahl (2005). Electronic patient–provider communication: will it offset office visits and telephone consultations in primary care? *International Journal of Medical Informatics 74*(9), 705–710.

Berkman, L. F. (1988). The changing and heterogeneous nature of aging and longevity: a social and biomedical perspective. In G. Maddox and M. Powell Lawton (Eds.), *Annual Review of Gerontology and Geriatrics*, pp. 37–68. New York, NY: Springer.

Berkman, L. F., T. Glass, I. Brissette, and T. E. Seeman (2000). From social integration to health: Durkheim in the new millennium. *Social Science & Medicine 51*(6), 843–857.

Berkman, N. D., S. Sheridan, K. E. Donahue, D. J. Halpern, and K. Crotty (2011). Low health literacy and health outcomes: An updated systematic review. *Annals of Internal Medicine 155*(2), 97–107.

Bernard, P., R. Charafeddine, K. L. Frohlich, M. Daniel, Y. Kestens, and L. Potvin (2007). Health inequalities and place: a theoretical conception of neighbourhood. *Social Science & Medicine 65*(9), 1839–1852.

Bessière, K., S. Kiesler, and R. Kraut (2008). Effects of Internet use and social resources on changes in depression. *Information, Communication & Society 11*(1), 47–70.

Bessière, K., S. Pressman, S. Kiesler, and R. Kraut (2010). Effects of Internet use on health and depression: a longitudinal study. *Journal of Medical Internet Research 12*(1), e6.

Birkin, M. and G. Clarke (2012). The enhancement of spatial microsimulation models using geodemographics. *The Annals of Regional Science 49*(2), 515–532.

Birkin, M. and M. Clarke (1988). SYNTHESIS – a synthetic spatial information system for urban and regional analysis: methods and examples. *Environment and Planning A 20*(12), 1645–1671.

Birru, M. S., V. M. Monaco, L. Charles, H. Drew, V. Njie, T. Bierria, E. Detlefsen, and R. A. Steinman (2004). Internet usage by low-literacy adults seeking health information: an observational analysis. *Journal of Medical Internet Research 6*(3), e25.

Blank, G. (2013). Who creates content? Stratification and content creation on the Internet. *Information, Communication & Society 16*(4), 590–612.

Blank, G. (2015). Age, income and education are the main predictors of Internet use and behaviour. In Oxford Internet Institute (Ed.), *The Oxford Internet Surveys: Internet Use, Behaviour and Attitudes in Great Britain 2003–2015*. Oxford: University of Oxford.

Blank, G., M. Graham, and C. Calvino (2014). Mapping the local geographies of digital inequality. Unpublished work.

Blank, G. and D. Grošelj (2014). Dimensions of Internet use: amount, variety, and types. *Information, Communication & Society 17*(4), 417–435.

Blaxter, M. (1990). *Health and lifestyles*. London: Tavistock/Routledge.

Blaxter, M. (2010). *Health*. Cambridge: Polity Press.

Blumler, J. G. (1979). The role of theory in uses and gratifications studies. *Communication Research 6*(1), 9–36.

Bollen, K. A. and J. Pearl (2013). Eight myths about causality and structural equation models. In *Handbook of Causal Analysis for Social Research*, pp. 301–328. Dordrecht: Springer.

Bottero, W. (2004). *Stratification: Social division and inequality*. Abingdon: Routledge.

Bourdieu, P. (1977). *Outline of a Theory of Practice*, Volume 16. Cambridge: Cambridge University Press.

Bourdieu, P. (1984). *Distinction: a social critique of the judgement of taste*. London: Routledge.

Bourdieu, P. (1986). The forms of capital. In A. H. Halsey, H. Lauder, P. Brown, and A. S. Wells (Eds.), *Education: Culture, Economy, and Society (1997)* (1 ed.)., pp. 46–58. Oxford: Oxford University Press.

Bourdieu, P. (1988). Vive la crise! *Theory and Society 17*(5), 773–787.

Bourdieu, P. (1990). *The Logic of Practice*. Stanford, CA: Stanford University Press.

Bowen, N. K. and S. Guo (2011). *Structural equation modeling*. Oxford: Oxford University Press.

Bowling, A. (2005). Just one question: if one question works, why ask several? *Journal of Epidemiology and Community Health 59*(5), 342–345.

Bowling, A. (2009). *Research Methods in Health: Investigating Health and Health Services* (3rd ed.). Maidenhead: Open University Press.

Branch, L., A. Jette, C. Evashwick, M. Polansky, G. Rowe, and P. Diehr (1981). Toward understanding elders' health service utilization. *Journal of Community Health 7*(2), 80–92.

Brewer, J. and A. Hunter (1989). *Multimethod Research: a Synthesis of Styles* (1 ed.). Newbury Park, CA; London: SAGE Publications.

British Academy (2014). *"If you could do one thing..." Nine local actions to reduce health inequalities*. London: The British Academy.

Brodie, M., R. E. Flournoy, D. E. Altman, R. J. Blendon, J. M. Benson, and M. D. Rosenbaum (2000). Health information, the Internet, and the digital divide. *Health Affairs 19*(6), 255–265.

Broom, A. (2005). Virtually he@lthy: the impact of Internet use on disease experience and the doctor-patient relationship. *Qualitative Health Research 15*(3), 325–345.

Brown, I., L. Brown, and D. Korff (2010). Using NHS patient data for research without consent. *Law, Innovation and Technology 2*(2), 219–258.

Brown, J. D. and K. Walsh-Childers (2009). Effects of media on personal and public health. In J. Bryant and D. Zillmann (Eds.), *Media Effects: Advances in Theory and Research* (3 ed.). New York, NY; London: Routledge.

Browne, M. W. and R. Cudeck (1992). Alternative ways of assessing model fit. *Sociological Methods & Research 21*(2), 230–258.

Browning, C. R. and K. A. Cagney (2002). Neighborhood structural disadvantage, collective efficacy, and self-rated physical health in an urban setting. *Journal of Health and Social Behavior 43*(4), 383–399.

Bryman, A. (2008). *Social Research Methods* (3rd ed.). Oxford: Oxford University Press.

BSA (2002). British Sociological Association. Statement of Ethical Practice for the British Sociological Association (March 2002). Accessed 25 January 2014, http://www.britsoc.co.uk/media/27107/StatementofEthicalPractice.pdf.

Bundorf, M. K., T. H. Wagner, and S. J. Singer (2006). Who searches the Internet for health information? *Health Services Research 41*(3p1), 819–836.

Buunk, B. P., F. X. Gibbons, and A. Buunk (2013). *Health, coping, and well-being: perspectives from social comparison theory.* Mahwah, NJ: Lawrence Erlbaum Associates.

Cabin, R. J. and R. J. Mitchell (2000). To Bonferroni or not to Bonferroni: when and how are the questions. *Bulletin of the Ecological Society of America 81*(3), 246–248.

Cacioppo, J. T., L. C. Hawkley, L. E. Crawford, J. M. Ernst, M. H. Burleson, R. B. Kowalewski, W. B. Malarkey, E. van Cauter, and G. G. Berntson (2002). Loneliness and health: potential mechanisms. *Psychosomatic Medicine 64*(3), 407–417.

Caiata-Zufferey, M., A. Abraham, K. Sommerhalder, and P. J. Schulz (2010). Online health information seeking in the context of the medical consultation in Switzerland. *Qualitative Health Research 20*(8), 1050–1061.

Campbell, C. (2001). Social capital and health: contextualising health promotion within local community networks. In S. Baron and J. Field (Eds.), *Social capital: critical perspectives* (1 ed.)., pp. 182–196. Oxford: Oxford University Press.

Carver, R. P. (1993). The case against statistical significance testing, revisited. *The Journal of Experimental Education 61*(4), 287–292.

Casiano, H., D. J. Kinley, L. Y. Katz, M. J. Chartier, and J. Sareen (2012). Media use and health outcomes in adolescents: findings from a nationally representative survey. *Journal of the Canadian Academy of Child and Adolescent Psychiatry 21*(4), 296–301.

Cassel, J. (1976). The contribution of the social environment to host resistance. *American Journal of Epidemiology 104*(2), 107–123.

Cassells, R., R. Miranti, and A. Harding (2013). Building a static spatial microsimulation model: data preparation. In R. Tanton and K. L. Edwards (Eds.), *Spatial Microsimulation: A Reference Guide for Users*, pp. 9–16. Dordrecht: Springer.

Cattell, V. (2001). Poor people, poor places, and poor health: the mediating role of social networks and social capital. *Social Science & Medicine 52*(10), 1501–1516.

Census (2011). Census for England and Wales. machine-readable single files obtained from the Office for National Statistics (ONS). Accessed 20 March – 14 May 2014, https://www.nomisweb.co.uk/census/2011/data_finder.

Census (2011a). Highest level of qualification by age (LC5102EW). Machine-readable dataset of the Office for National Statistics. Accessed 13 May 2014, https://www.nomisweb.co.uk/census/2011/lc5102ew.

Census (2011b). Long-term health problem or disability by general health by sex by age (LC3302EW). Machine-readable dataset of the Office for National Statistics. Accessed 13 May 2014, https://www.nomisweb.co.uk/census/2011/lc3302ew.

Census (2011c). General health by highest level of qualification (CT0465). Machine-readable dataset of the Office for National Statistics (published ad-hoc data and analysis). Accessed 25 June 2015, http://www.ons.gov.uk/ons/about-ons/business-transparency/freedom-of-information/what-can-i-request/published-ad-hoc-data/census/qualifications/ct0465-2011-census.zip.

Census (2011d). General health by NS-SeC (LC3601EW). Machine-readable dataset of the Office for National Statistics. Accessed 27 November 2014, http://www.nomisweb.co.uk/census/2011/lc3601ew.

Census (2011e). OA population weighted centroids. Machine-readable dataset of the Office for National Statistics. Accessed 15 May 2014, http://www.ons.gov.uk/ons/external-links/social-media/g-m/2011-oa-population-weighted-centroids.html.

Census (2011f). NS-SeC by sex (LC6121EW). Machine-readable dataset of the Office for National Statistics. Accessed 13 May 2014, http://www.nomisweb.co.uk/census/2011/lc6121ew.

Census (2011g). Age by single year (QS103EW). Machine-readable dataset of the Office for National Statistics. Accessed 13 May 2014, https://www.nomisweb.co.uk/census/2011/qs103ew.

Census (2011h). Highest level of qualification by ethnic group (LC5202EW). Machine-readable dataset of the Office for National Statistics. Accessed 13 May 2014, https://www.nomisweb.co.uk/census/2011/lc5202ew.

Census (2011i). Sex by age (LC1117EW). Machine-readable dataset of the Office for National Statistics. Accessed 14 December 2015, http://www.nomisweb.co.uk/census/2011/lc1117ew.

Census (2011j). Rural-urban classification (2011) of output areas (2011) E+W. Machine-readable dataset of the Office for National Statistics. Accessed 20 November 2013, https://geoportal.statistics.gov.uk/Docs/Products/Rural-urban_classification_(2011)_of_output_areas_(2011)_E+W.zip.

Census (2011k). 2011 OAC Clusters and Names. Machine-readable dataset of the Office for National Statistics. Accessed 22 August 2014, http://www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/ns-area-classifications/ns-2011-area-classifications/datasets/2011-oac-clusters-and-names-csv.zip.

Census (2011l). Output areas (2011) to regions (2011) lookup. Machine-readable dataset of the Office for National Statistics. Accessed 22 August 2014, https://geoportal.statistics.gov.uk/Docs/Lookups/Output_areas_(2011)_to_regions_(2011)_Eng_lookup.zip.

Census (2011m). Output areas (2001) to output areas (2011) to local authority districts (2011) lookup. Machine-readable dataset of the Office for National Statistics. Accessed 13 May 2014, https://geoportal.statistics.gov.uk/Docs/Lookups/Output_areas_(2001)_to_output_areas_(2011)_to_local_authority_districts_(2011)_E+W_lookup.zip.

Census (2011n). Output areas (2011) to wards (2013) to local authority districts (2013) lookup. Machine-readable dataset of the Office for National Statistics. Accessed 20 August 2015, https://geoportal.statistics.gov.uk/Docs/Lookups/Output_areas_(2011)_to_wards_(2013)_to_local_authority_districts_(2013)_E+W_lookup.zip.

Chandola, T. and A. Jenkins (2014). The scope of adult and further education for reducing health inequalities. In The British Academy (Ed.), *"If you could do one thing..." Nine local actions to reduce health inequalities*, pp. 82–90. London: The British Academy.

Chapple, A., J. Evans, and S. Ziebland (2012). An alarming prognosis: how people affected by pancreatic cancer use (and avoid) internet information. *Policy & Internet 4*(2), 1–20.

Chaudhry, B., J. Wang, S. Wu, M. Maglione, W. Mojica, E. Roth, S. C. Morton, and P. G. Shekelle (2006). Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Annals of Internal Medicine 144*(10), 742–752.

Chin, W. W. and P. A. Todd (1995). On the use, usefulness, and ease of use of structural equation modeling in mis research: a note of caution. *MIS quarterly*, 237–246.

Cho, J., H. G. De Zúñiga, H. Rojas, and D. V. Shah (2003). Beyond access: the digital divide and Internet uses and gratifications. *IT & Society 1*(4), 46–72.

Christodoulides, G., N. Michaelidou, and N. T. Siamagka (2013). A typology of internet users based on comparative affective states: evidence from eight countries. *European Journal of Marketing 47*(1/2), 153–173.

Clarke, G. and M. Madden (2001). *Regional Science in Business*. Berlin; Heidelberg: Springer Science & Business Media.

Clatworthy, J., D. Buick, M. Hankins, J. Weinman, and R. Horne (2005). The use and reporting of cluster analysis in health psychology: a review. *British Journal of Health Psychology 10*(3), 329–358.

Cline, R. J. W. and K. M. Haynes (2001). Consumer health information seeking on the Internet: the state of the art. *Health Education Research 16*(6), 671–692.

Cobb, S. (1976). Social support as a moderator of life stress. *Psychosomatic Medicine 38*(5), 300–314.

Cockerham, W. C. (2005). Health lifestyle theory and the convergence of agency and structure. *Journal of Health and Social Behavior 46*(1), 51–67.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist 45*(12), 1304.

Cohen, J. (1992). A power primer. *Psychological Bulletin 112*(1), 155.

Cohen, N. and T. Arieli (2011). Field research in conflict environments: methodological challenges and snowball sampling. *Journal of Peace Research 48*(4), 423–435.

Cohen, S. (1988). Psychosocial models of the role of social support in the etiology of physical disease. *Health Psychology 7*(3), 269–297.

Cohen, S. and G. McKay (1984). Social support, stress and the buffering hypothesis: a theoretical analysis. In A. Baum, S. E. Taylor, and J. E. Singer (Eds.), *Handbook of Psychology and Health*, pp. 253–267. Hillsdale, NJ; London: Erlbaum.

Cohen, S. and T. A. Wills (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin 98*(2), 310.

Connelly, R., C. J. Playford, V. Gayle, and C. Dibben (2016). The role of administrative data in the big data revolution in social science research. *Social Science Research*, 1–12.

Conole, G. and M. Dyke (2004). What are the affordances of information and communication technologies? *Association for Learning Technology Journal 12*(2), 113–124.

Cotten, S. R. and S. S. Gupta (2004). Characteristics of online and offline health information seekers and factors that discriminate between them. *Social Science & Medicine 59*(9), 1795–1806.

Coulter, A. (2011). *Engaging patients in healthcare*. Maidenhead: Open University Press.

Coyne, I. T. (1997). Sampling in qualitative research. purposeful and theoretical sampling; merging or clear boundaries? *Journal of Advanced Nursing 26*(3), 623–630.

Crawford, R. and C. Emmerson (2012). NHS and social care funding: the outlook to 2021/22. Accessed 05 March 2015, http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/120704_nhs-social-care-funding-outlook-2021-22-update2.pdf.

Creswell, J. W. and V. L. P. Clark (2011). *Designing and Conducting Mixed Methods Research* (2nd ed.). London: SAGE.

Creswell, J. W. and D. L. Miller (2000). Determining validity in qualitative inquiry. *Theory into practice 39*(3), 124–130.

Crocco, A. G., M. Villasis-Keever, and A. R. Jadad (2002). Analysis of cases of harm associated with use of health information on the Internet. *Journal of the American Medical Association 287*(21), 2869–2871.

Curtis, S. and I. Rees Jones (1998). Is there a place for geography in the analysis of health inequality? *Sociology of Health & Illness 20*(5), 645–672.

Dark, S. and D. Bram (2007). The modifiable areal unit problem (MAUP) in physical geography. *Progress in Physical Geography 31*(5), 471–479.

Davies, T. (2013). Open Data Barometer. 2013 Global Report, World Wide Web Foundation & Open Data Institute. Accessed 11 February 2014, http://www.opendataresearch.org/dl/odb2013/Open-Data-Barometer-2013-Global-Report.pdf.

Deetjen, U. (2016a). The lifestyle paradox: how online health information adversely influences health perception. *Paper submitted to journal* (forthcoming).

Deetjen, U. (2016b). European e-prescriptions: benefits and success factors. Working Paper No. 5, Working Paper Series of the Cyber Studies Programme, Department of International Relations, University of Oxford.

Deetjen, U., E. T. Meyer, and R. Schroeder (2015). Big data for advancing dementia research: an evaluation of data sharing practices in research on age-related neurodegenerative diseases. OECD Digital Economy Papers no. 246, Paris, France: OECD Publishing.

Deetjen, U. and J. A. Powell (2016a). Internet use and health: connecting secondary data through spatial microsimulation. *Paper submitted to journal* (forthcoming).

Deetjen, U. and J. A. Powell (2016b). Internet use and the outcomes digital divide: a typology of health information seekers. *paper submitted to journal* (forthcoming).

Deetjen, U. and J. A. Powell (2016c). How does the Internet affect health service use? A new theoretical model and mixed methods study. *Paper submitted to journal* (forthcoming).

Deetjen, U. and J. A. Powell (2016d). Informational and Emotional Elements in Online Support Groups: a Bayesian Approach to Large-scale Content Analysis. *Journal of the American Informatics Association*.

Deming, W. E. and F. F. Stephan (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics 11*(4), 427–444.

Dening, T. R., L.-Y. Chi, C. Brayne, F. A. Huppert, E. S. Paykel, and D. W. O'Connor (1998). Changes in self-rated health, disability and contact with services in a very elderly cohort: a 6-year follow-up study. *Age and Ageing 27*(1), 23–33.

Denzin, N. K. and Y. S. Lincoln (1994). *Handbook of qualitative research*. Thousand Oaks, CA: SAGE.

Department of Health (1999). Saving Lives: Our Healthier Nation. Technical report.

Department of Health (2001). The Expert Patient: A New Approach to Chronic Disease Management for the 21st Century. Technical report.

Department of Health (2010). Equity and Excellence: Liberating the NHS. Technical report.

Department of Health (2011). Liberating the NHS: An Information Revolution. Technical report.

Department of Health (2012a). Liberating the NHS: Greater Choice and Control. Choice of named consultant-led team Equality Analysis. Technical report.

Department of Health (2012b). Liberating the NHS: No decision about me, without me – Government response to the consultation. Technical report.

Department of Health (2012c). The Power of Information: Putting all of us in control of the health and care information we need. Technical report.

Department of Health (2012d). Long Term Conditions Compendium of Information. Technical report.

Department of Health (2013). Department of Health Corporate Plan 2013–2014. Accessed 06 February 2014, https://www.gov.uk/government/publications/department-of-health-corporate-plan-2013-14/department-of-health-2013-14-corporate-plan.

DeWalt, D. A., N. D. Berkman, S. Sheridan, K. N. Lohr, and M. P. Pignone (2004). Literacy and health outcomes. *Journal of General Internal Medicine 19*(12), 1228–1239.

Dey, I. (2005). *Qualitative data analysis: A user friendly guide for social scientists*. London; New York, NY: Routledge.

Di Iorio, C., F. Carinci, and J. Oderkirk (2014). Health research and systems' governance are at risk: should the right to data protection override health? *Journal of Medical Ethics 40*(7), 488–492.

DiCicco-Bloom, B. and B. F. Crabtree (2006). The qualitative research interview. *Medical Education 40*(4), 314–321.

Dietz, R. D. (2002). The estimation of neighborhood effects in the social sciences: an interdisciplinary approach. *Social Science Research 31*(4), 539–575.

Dillman, D. A., M. D. Sinclair, and J. R. Clark (1993). Effects of questionnaire length, respondent-friendly design, and a difficult question on response rates for occupant-addressed census mail surveys. *Public Opinion Quarterly 57*(3), 289–304.

Dillon, B. (2009). *Tormented hope: nine hypochondriac lives*. Dublin: Penguin.

DiMaggio, P. and E. Hargittai (2001). From the "digital divide" to "digital inequality": studying Internet use as penetration increases. Working paper series, Princeton University Center for Arts and Cultural Policy Studies.

Dingwall, R. (2001). *Aspects of illness*. Aldershot: Ashgate Pub Limited.

Dobransky, K. and E. Hargittai (2006). The disability divide in Internet access and use. *Information, Communication & Society 9*(3), 313–334.

Dorling, D., G. Smith, M. Noble, G. Wright, R. Burrows, J. Bradshaw, H. Joshi, C. Pattie, R. Mitchell, A. E. Green, and A. McCulloch (2001). How much does place matter? *Environment and Planning A 33*(8), 1335–1369.

Dörner, K. and D. Edelman (2015). What "digital" really means. *McKinsey Digital July 2015*.

Doty, D. H. and W. H. Glick (1994). Typologies as a unique form of theory building: toward improved understanding and modeling. *Academy of Management Review 19*(2), 230–251.

Due, P., B. Holstein, R. Lund, J. Modvig, and K. Avlund (1999). Social relations: network, support and relational strain. *Social Science & Medicine 48*(5), 661–673.

Duncan, O. D. and B. Davis (1953). An alternative to ecological correlation. *American Sociological Review 18*(6), 665–666.

Duncan, T. E. and E. McAuley (1993). Social support and efficacy cognitions in exercise adherence: a latent growth curve analysis. *Journal of Behavioral Medicine 16*(2), 199–218.

Durkheim, E. (1952). *Suicide: a Study in Sociology*. London: Routledge & Kegan Paul Lt. Edited by G. Simpson, translated by J. A. Spaulding and G. Simpson.

Dutta-Bergman, M. J. (2006). Media use theory and Internet use for health care. In M. Murero and R. E. Rice (Eds.), *The Internet and Health Care: Theory, Research, and Practice*, pp. 83–103. Mahwah, NJ: Lawrence Erlbaum Associates.

Dutton, W. H. and G. Blank (2011). Next Generation Users: the Internet in Britain. Oxford Internet Survey 2011, Oxford Internet Institute, University of Oxford.

Dutton, W. H. and G. Blank, with assistance by D. Grošelj (2013). Cultures of the Internet: the Internet in Britain. Oxford Internet Survey 2013, Oxford Internet Institute, University of Oxford.

Dutton, W. H. and A. Shepherd (2006). Trust in the Internet as an experience technology. *Information, Communication & Society 9*(4), 433–451.

Dwyer, D. S. and H. Liu (2013). The impact of consumer health information on the demand for health services. *The Quarterly Review of Economics and Finance 53*(1), 1–11.

Eastin, M. S. and R. LaRose (2000). Internet self-efficacy and the psychology of the digital divide. *Journal of Computer-Mediated Communication 6*(1), 0.

Edwards, K. L. and G. Clarke (2013). SimObesity: combinatorial optimisation (deterministic) model. In R. Tanton and K. L. Edwards (Eds.), *Spatial Microsimulation: A Reference Guide for Users*, pp. 69–85. Dordrecht: Springer.

Edwards, K. L., G. P. Clarke, J. Thomas, and D. Forman (2011). Internal and external validation of spatial microsimulation models: small area estimates of adult obesity. *Applied Spatial Analysis and Policy 4*(4), 281–300.

Edwards, K. L. and R. Tanton (2013). Validation of Spatial Microsimulation Models. In R. Tanton and K. L. Edwards (Eds.), *Spatial Microsimulation: A Reference Guide for Users*, pp. 69–85. Dordrecht: Springer.

El Emam, K., S. Rodgers, and B. Malin (2015). Anonymising and sharing individual patient data. *British Medical Journal 350*, h1139.

Entwistle, V. A., E. F. France, S. Wyke, R. Jepson, K. Hunt, S. Ziebland, and A. Thompson (2011). How information about other people's personal experiences can help with healthcare decision-making: a qualitative study. *Patient Education and Counseling 85*(3), e291–e298.

Ericsson, Å. (1997). The importance of lifestyle to self-assessed health. *Health Policy 42*(2), 145–155.

ESRC (2009). Economic and Social Research Council. Strategic plan 2009–2014: Delivering impact through social science. Accessed 16 April 2014, http://www.esrc.ac.uk/Image/Strategic_Plan_FINAL_tcm11-13164.pdf.

ESRC (2015a). Economic and Social Research Council. Strategic plan 2015: Delivering impact through social science. Accessed 5 October 2015, http://www.esrc.ac.uk/files/news-events-and-publications/publications/corporate-publications/strategic-plan/esrc-strategic-plan-2015.

ESRC (2015b). Economic and Social Research Council. Framework for Research Ethics (FRE), updated January 2015. Accessed 10 March 2015, http://www.esrc.ac.uk/files/funding/guidance-for-applicants/esrc-framework-for-research-ethics-2015.

European Commission (2012). eHealth Action Plan 2012-2020. Accessed 06 February 2014, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=4188.

European Commission (2013). Digital Agenda Scoreboard 2013. Accessed 06 February 2014, https://ec.europa.eu/digital-agenda/sites/digital-agenda/files/DAESCOREBOARD2013-SWD2013217FINAL.pdf.

European Commission (2015). Major and chronic diseases (MCDs). Accessed 28 February 2015, http://ec.europa.eu/health/major_chronic_diseases/diseases/index_en.htm.

Everitt, B. S., S. Landau, M. Leese, and D. Stahl (2011). *Cluster Analysis* (5 ed.). Chichester: Wiley.

Eynon, R., U. Deetjen, and L. E. Malmberg (2016). Moving up in the information society? The relationship between Internet use and social mobility. *Paper currently under preparation* (forthcoming).

Eynon, R., J. Fry, and R. Schroeder (2008). The Ethics of Internet Research. In N. G. Fielding, R. M. Lee, and G. Blank (Eds.), *The SAGE Handbook of Online Research Methods*, pp. 23–41. London: SAGE.

Eynon, R. and L.-E. Malmberg (2012). Understanding the online information-seeking behaviours of young people: the role of networks of support. *Journal of Computer Assisted Learning 28*(6), 514–529.

Eysenbach, G. (2003). The impact of the Internet on cancer outcomes. *CA Cancer Journal for Clinicians 53*(6), 356–371.

Eysenbach, G. (2008). Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. *Journal of Medical Internet Research 10*(3), e22.

Eysenbach, G., J. Powell, M. Englesakis, C. Rizo, and A. Stern (2004). Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *British Medical Journal 328*(7449), 1166.

Eysenbach, G., J. Powell, O. Kuss, and E.-R. Sa (2002). Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *Journal of the American Medical Association 287*(20), 2691–2700.

Feilzer, M. Y. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research 4*(1), 6–16.

Fergus, T. A. (2013). Cyberchondria and intolerance of uncertainty: examining when individuals experience health anxiety in response to Internet searches for medical information. *Cyberpsychology, Behavior, and Social Networking 16*(10), 735–739.

Ferlander, S. (2007). The importance of different forms of social capital for health. *Acta Sociologica 50*(2), 115–128.

Ferraro, K. F. (1980). Self-ratings of health among the old and the old-old. *Journal of Health and Social Behavior 21*(4), 377–383.

Ferreira, L. and D. B. Hitchcock (2009). A comparison of hierarchical methods for clustering functional data. *Communications in Statistics-Simulation and Computation 38*(9), 1925–1949.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations 7*(2), 117–140.

Fienberg, S. E. (1970). An iterative procedure for estimation in contingency tables. *The Annals of Mathematical Statistics 41*(3), 907–917.

Fillenbaum, G. G. (1979). Social context and self-assessments of health among the elderly. *Journal of Health and Social Behavior 20*(1), 45–51.

Fiscella, K., P. Franks, and C. M. Clancy (1998). Skepticism toward medical care and health care utilization. *Medical Care 36*(2), 180–189.

Fogel, J., S. M. Albert, F. Schnabel, and B. A. Ditkoff (2002). Internet use and social support in women with breast cancer. *Health Psychology 21*(4), 398–404.

Folland, S., A. C. Goodman, and M. Stano (2014). *The economics of health and health care*, Volume 7. Harlow: Pearson.

Forkner-Dunn, J. (2003). Internet-based patient self-care: the next generation of health care delivery. *Journal of Medical Internet Research 5*(2), e8.

Fowler, F. J. (2014). *Survey research methods* (5 ed.). Thousand Oaks, CA: SAGE publications.

Fox, S. (2011). *The social life of health information.* Washington, DC: Pew Internet & American Life Project.

Fox, S. and L. Rainie (2002). Vital decisions: How Internet users decide what information to trust when they or their loved ones are sick. Technical report, Pew Internet and American Life Project, Washington D.C.

France, E. F., S. Wyke, S. Ziebland, V. A. Entwistle, and K. Hunt (2011). How personal experiences feature in women's accounts of use of information for decisions about antenatal diagnostic testing for foetal abnormality. *Social Science & Medicine 72*(5), 755–762.

Fratiglioni, L., H.-X. Wang, K. Ericsson, M. Maytan, and B. Winblad (2000). Influence of social network on occurrence of dementia: a community-based longitudinal study. *The Lancet 355*(9212), 1315–1319.

Freedman, D. A. (1999). Ecological inference and the ecological fallacy. *International Encyclopedia of the Social & Behavioral Sciences 6*, 4027–4030.

Freese, J. (2009). Secondary analysis of large social surveys. In E. Hargittai (Ed.), *Research confidential: Solutions to problems most social scientists pretend they never have*, pp. 238–261. Ann Arbor, MI: University of Michigan Press.

Fuchs, C. (2008). *Internet and society: social theory in the information age.* New York; London: Routledge.

Gan, G., C. Ma, and J. Wu (2007). *Data clustering: theory, algorithms, and applications*, Volume 20 of *ASA-SIAM Series on Statistics and Applied Probability.* Philadelphia, PA: Society for Industrial and Applied Mathematics and Alexandria; VA: American Statistical Association.

Garamszegi, L. Z. (2006). Comparing effect sizes across variables: generalization without the need for bonferroni correction. *Behavioral Ecology 17*(4), 682–687.

Garrity, T. F., G. W. Somes, and M. B. Marx (1978). Factors influencing self-assessment of health. *Social Science & Medicine. Part A: Medical Psychology & Medical Sociology 12*, 77–81.

Gately, C., A. Rogers, and C. Sanders (2007). Re-thinking the relationship between long-term condition self-management education and the utilisation of health services. *Social Science & Medicine 65*(5), 934–945.

Gaver, W. W. (1991). Technology affordances. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 79–84. ACM.

Gefen, D., D. Straub, and M.-C. Boudreau (2000). Structural equation modeling and regression: guidelines for research practice. *Communications of the Association for Information Systems 4*(1), 7.

Geronimus, A. T. and J. Bound (1998). Use of census-based aggregate variables to proxy for socioeconomic group: evidence from national samples. *American Journal of Epidemiology 148*(5), 475–486.

Gershuny, J. (1983). *Social innovation and the division of labour*. Oxford: Oxford University Press.

Gibson, J. J. (1986). *The ecological approach to visual perception*. Hillsdale, NJ; London: Lawrence Erlbaum Associates.

Giddens, A. (1984). *The Constitution of Society: Outline of the Theory of Structuration*. Cambridge: Polity.

Giddens, A. (1991). *Modernity and self-identity: self and society in the late modern age*. Stanford, CA: Stanford University Press.

Gill, P., K. Stewart, E. Treasure, and B. Chadwick (2008). Methods of data collection in qualitative research: interviews and focus groups. *British Dental Journal 204*(6), 291–295.

Gilmour, J. A. (2007). Reducing disparities in the access and use of Internet health information. a discussion paper. *International Journal of Nursing Studies 44*(7), 1270–1278.

Glogowska, M. (2011). Paradigms, pragmatism and possibilities: mixed-methods research in speech and language therapy. *International Journal of Language & Communication Disorders 46*(3), 251–260.

Gneezy, U., S. Meier, and P. Rey-Biel (2011). When and why incentives (don't) work to modify behavior. *The Journal of Economic Perspectives 25*(4), 191–209.

Go ON UK (2015). Digital Exclusion Heatmap: Exploring exclusion from a digital United Kingdom. Accessed 21 October 2015, http://www.go-on.co.uk/resources/heatmap.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Garden City, NY: Doubleday.

Goldsmith, H. F., D. J. Jackson, and R. L. Hough (1988). Process model of seeking mental health services: proposed framework for organizing the literature on help-seeking. In *Needs Assessment: Its future*, pp. 49–64. Rockville, MD: US Department of Health and Human Services, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, Nation Institute of Mental Health, Division of Biometry and Applied Sciences.

Goldstein, H. (1999). Multilevel statistical models.

Goldstein, M. S., J. M. Siegel, and R. Boyer (1984). Predicting changes in perceived health status. *American Journal of Public Health 74*(6), 611–614.

Goodman, L. A. (1953). Ecological regressions and behavior of individuals. *American Sociological Review 18*(6), 663–664.

Goodwin, R. and G. Engstrom (2002). Personality and the perception of health in the general population. *Psychological medicine 32*(02), 325–332.

Gray, D. E. (2009). *Doing Research in the Real World* (2 ed.). London: SAGE Publications Ltd.

Gray, M. J. A. (2002). *The resourceful patient*. Osney Mead, Oxford: Alden Group.

Green, M. and P. Rossall (2013). Digital Inclusion Evidence Review. Technical report, Age UK.

Greene, J. C., V. J. Caracelli, and W. F. Graham (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis 11*(3), 255–274.

Greene, W. H. (2003). *Econometric analysis*. Pearson Education India.

Greener, I. (2008). Expert patients and human agency: long-term conditions and giddens' structuration theory. *Social Theory & Health 6*(4), 273–290.

Grembowski, D., D. Patrick, P. Diehr, M. Durham, S. Beresford, E. Kay, and J. Hecht (1993). Self-efficacy and health behavior among older adults. *Journal of Health and Social Behavior 34*(2), 89–104.

Grossman, M. (1972). On the concept of health capital and the demand for health. *The Journal of Political Economy 80*(2), 223–255.

Gustafson, D. H., R. Hawkins, E. Boberg, S. Pingree, R. E. Serlin, F. Graziano, and C. L. Chan (1999). Impact of a patient-centered, computer-based health information/support system. *American Journal of Preventive Medicine 16*(1), 1–9.

Gustafson, D. H., R. Hawkins, S. Pingree, F. McTavish, N. K. Arora, J. Mendenhall, D. F. Cella, R. C. Serlin, F. M. Apantaku, J. Stewart, et al. (2001). Effect of computer support on younger women with breast cancer. *Journal of General Internal Medicine 16*(7), 435–445.

Hale, T. M. (2013). Is there such a thing as an online health lifestyle? Examining the relationship between social status, Internet access, and health behaviors. *Information, Communication & Society 16*(4), 501–518.

Hale, T. M., S. R. Cotten, P. Drentea, and M. Goldner (2010). Rural–urban differences in general and health-related Internet use. *American Behavioral Scientist 53*(9), 1304–1325.

Hall, M. A., E. Dugan, B. Zheng, and A. K. Mishra (2001). Trust in physicians and medical institutions: what is it, can it be measured, and does it matter? *Milbank Quarterly 79*(4), 613–639.

Ham, C. (2009). *Health Policy in Britain* (6 ed.). Public Policy and Politics. Basingstoke: Palgrave Macmillan.

Hampton, K. and B. Wellman (2003). Neighboring in Netville: how the Internet supports community and social capital in a wired suburb. *City and Community 2*(4), 277–311.

Hannay, D. (1980). The "iceberg" of illness and "trivial" consultations. *The Journal of the Royal College of General Practitioners 30*(218), 551–554.

Hannay, D. R. (1979). *The symptom iceberg: a study of community health*. London; Boston: Routledge & Kegan Paul.

Hansen, M. S., P. Fink, M. Frydenberg, M.-L. Oxhøj, L. Søndergaard, and M. Eriksen (2001). Mental disorders in medical inpatients and the association to severity of illness, self-rated physical disability, and health perception. *Psychosomatics 42*(1), 41–47.

Hardey, M. (1999). Doctor in the house: the Internet as a source of lay health knowledge and the challenge to expertise. *Sociology of Health & Illness 21*(6), 820–835.

Hargittai, E. (2002). Second-level digital divide: differences in people's online skills. *First Monday 7*(4).

Hargittai, E. and S. Shafer (2006). Differences in actual and perceived online skills: the role of gender. *Social Science Quarterly 87*(2), 432–448.

Haslett, S., G. Jones, A. Noble, and D. Ballas (2010). More for less? Comparing small area estimation, spatial microsimulation, and mass imputation. In *Proceedings of the Survey Research Methods Section*, pp. 1584–1598.

Haythornthwaite, C. and B. Wellman (2002). Moving the Internet out of cyberspace. The Internet in everyday life: an introduction. In B. Wellman and C. Haythornthwaite (Eds.), *The Internet in Everyday Life*, Information Age Series. Oxford: Blackwell.

Helsper, E. (2008). Digital inclusion: an analysis of social disadvantage and the information society. Technical report, Department for Communities and Local Government, London, UK.

Helsper, E. J. (2012). A corresponding fields model for the links between social and digital exclusion. *Communication Theory 22*(4), 403–426.

Henwood, F., S. Wyatt, A. Hart, and J. Smith (2003). 'Ignorance is bliss sometimes': constraints on the emergence of the 'informed patient' in the changing landscapes of health information. *Sociology of Health & Illness 25*(6), 589–607.

Hermes, K. and M. Poulsen (2012). A review of current methods to generate synthetic spatial microdata using reweighting and future directions. *Computers 36*(4), 281–290.

Herod, A. (1993). Gender issues in the use of interviewing as a research method. *The Professional Geographer 45*(3), 305–317.

HES (2013). Hospital Episode Statistics. Machine-readable dataset obtained from Health and Social Care Information Centre (HSCIC). Reference No. NIC-317612-V1Q3X.

Hoffman, S. (2010). Electronic health records and research: privacy versus scientific priorities. *The American Journal of Bioethics 10*(9), 19–20.

Hoffmann, C. P., C. Lutz, and M. Meckel (2015). Content creation on the Internet: a social cognitive perspective on the participation divide. *Information, Communication & Society 18*(6), 696–716.

Holstein, J. A. and J. F. Gubrium (2004). The active interview. In D. Silverman (Ed.), *Qualitative research: Theory, method and practice* (2 ed.)., pp. 140–161. London; Thousand Oaks, CA: SAGE Publications.

Holt-Lunstad, J., T. B. Smith, and J. B. Layton (2010). Social relationships and mortality risk: a meta-analytic review. *PLoS Medicine 7*(7), e1000316.

Houston, T. K. and J. J. Allison (2002). Users of Internet health information: differences by health status. *Journal of Medical Internet Research 4*(2), e7.

Hox, J. (1998). Multilevel modeling: When and why. In *Classification, data analysis, and data highways*, pp. 147–154. Springer.

HSCIC (2014). Data Access Advisory Group, Minutes of meeting held 12 November 2014. Accessed 19 November 2014, not available online anymore.

HSCIC (2015). Admitted Patient Care, England, 2013-14. Technical report, Health and Social Care Information Centre.

Hughes, N., L. Locock, and S. Ziebland (2013). Personal identity and the role of 'carer' among relatives and friends of people with multiple sclerosis. *Social Science & Medicine 96*, 78–85.

Hunt, S. M., S. McKenna, J. McEwen, E. Backett, J. Williams, and E. Papp (1980). A quantitative approach to perceived health status: a validation study. *Journal of Epidemiology and Community Health 34*(4), 281–286.

Idler, E. L. and Y. Benyamini (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior 38*(1), 21–37.

Jain, A. K., M. N. Murty, and P. J. Flynn (1999). Data clustering: a review. *ACM computing surveys (CSUR) 31*(3), 264–323.

James, N. and H. Busher (2009). *Online Interviewing*. London: SAGE Publications.

Janesick, V. J. (1994). The dance of qualitative research design: metaphor, methodolatry, and meaning. In *Strategies of Qualitative Inquiry*, pp. 35–55. Thousand Oaks, CA; London: SAGE Publications.

Janz, N. K. and M. H. Becker (1984). The health belief model: a decade later. *Health Education & Behavior 11*(1), 1–47.

Jenkins, R. (2001). *Pierre Bourdieu* (2 ed.). Key Sociologists. London: Routledge.

Jennings, M. K. and R. G. Niemi (1975). Continuity and change in political orientations: a longitudinal study of two generations. *American Political Science Review 69*(4), 1316–1335.

Jha, A. K., D. Doolan, D. Grandt, T. Scott, and D. W. Bates (2008). The use of health information technology in seven nations. *International Journal of Medical Informatics 77*(12), 848–854.

Johnson, R. B. and A. J. Onwuegbuzie (2004). Mixed methods research: a research paradigm whose time has come. *Educational Researcher 33*(7), 14–26.

Johnson, R. J. and F. D. Wolinsky (1993). The structure of health status among older adults: disease, disability, functional limitation, and perceived health. *Journal of Health and Social Behavior 34*(2), 105–121.

Kane, R. L. and D. M. Radosevich (2011). *Conducting Health Outcomes Research*. Sudbury, MA: Jones & Bartlett Publishers.

Kaplan, G. A. and T. Camacho (1983). Perceived health and mortality: a nine-year follow-up of the human population laboratory cohort. *American Journal of Epidemiology 117*(3), 292–304.

Katz, E. (1959). Mass communications research and the study of popular culture: an editorial note on a possible future for this journal. *Studies in public communication 2*, 1.

Katz, E., J. G. Blumler, and M. Gurevitch (1973). Uses and gratifications research. *Public Opinion Quarterly 37*(4), 509–523.

Kawachi, I. and L. F. Berkman (2003). Introduction. In I. Kawachi and L. F. Berkman (Eds.), *Neighborhoods and Health*, pp. 1–19. Oxford: Oxford University Press.

Kawachi, I., B. P. Kennedy, and R. Glass (1999). Social capital and self-rated health: a contextual analysis. *American Journal of Public Health 89*(8), 1187–1193.

Kemper, E. A., S. Stringfield, and C. Teddlie (2003). Mixed methods sampling strategies in social science research. In *Handbook of Mixed Methods in Social and Behavioral Research*, pp. 273–296. Thousand Oaks, CA: SAGE.

King, G. (1997). *A solution to the ecological inference problem*. Princeton, NJ: Princeton University Press.

King, G., O. Rosen, and M. A. Tanner (2004). *Ecological inference: new methodological strategies*. Cambridge: Cambridge University Press.

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement 56*(5), 746–759.

Kitto, S. C., J. Chesters, and C. Grbich (2008). Quality in qualitative research. *Medical Journal of Australia 188*(4), 243.

Kivinen, P., P. Halonen, M. Eronen, and A. Nissinen (1998). Self-rated health, physician-rated health and associated factors among elderly men: the Finnish cohorts of the seven countries study. *Age and Ageing 27*(1), 41–47.

Kivits, J. (2009). Everyday health and the Internet: a mediated health perspective on health information seeking. *Sociology of Health & Illness 31*(5), 673–687.

Kleinman, A. (1978). Concepts and a model for the comparison of medical systems as cultural systems. *Social Science & Medicine. Part B: Medical Anthropology 12*, 85–93.

Knox, S. S. and K. Uvnäs-Moberg (1998). Social isolation and cardiovascular disease: an atherosclerotic pathway? *Psychoneuroendocrinology 23*(8), 877–890.

Koos, E. L. (1954). *The health of Regionville; What the people thought and did about it*. New York, NY: Columbia University Press.

Kraut, R., S. Kiesler, B. Boneva, and J. Cummings (2002). Internet paradox revisited. *Journal of Social Issues 58*(1), 49–74.

Kraut, R., M. Patterson, V. Lundmark, S. Kiesler, T. Mukopadhyay, and W. Scherlis (1998). Internet paradox: a social technology that reduces social involvement and psychological well-being? *The American Psychologist 53*(9), 1017–1031.

Kross, E., P. Verduyn, E. Demiralp, J. Park, D. S. Lee, N. Lin, H. Shablack, J. Jonides, and O. Ybarra (2013). Facebook use predicts declines in subjective well-being in young adults. *PloS one 8*(8), e69841.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity 47*(4), 2025–2047.

La Gaipa, J. J. (1990). The negative effects of information social support systems. In S. Duck and R. C. Silver (Eds.), *Personal Relationships and Social Support*, pp. 122–139. London: SAGE.

Lahiri, P. and M. D. Larsen (2005). Regression analysis with linked data. *Journal of the American Statistical Association 100*(469), 222–230.

LaRose, R. and M. S. Eastin (2004). A social cognitive theory of Internet uses and gratifications: toward a new model of media attendance. *Journal of Broadcasting & Electronic Media 48*(3), 358–377.

Lauckner, C. and G. Hsieh (2013). The presentation of health-related search results and its impact on negative emotional outcomes. In *Proceedings of the CHI 2013 Conference*, April 27–May 2, 2013, Paris, France.

LaVela, S. L., B. Smith, F. M. Weaver, and S. A. Miskevics (2004). Geographical proximity and health care utilization in veterans with SCI&D in the USA. *Social Science & Medicine 59*(11), 2387–2399.

Lee, C.-J. (2008). Does the Internet displace health professionals? *Journal of Health Communication 13*(5), 450–464.

Lee, C.-J. and R. C. Hornik (2009). Physician trust moderates the Internet use and physician visit relationship. *Journal of Health Communication 14*(1), 70–76.

Lee, J. K. (2007). The effect of the internet on homogeneity of the media agenda: A test of the fragmentation thesis. *Journalism & Mass Communication Quarterly 84*(4), 745–760.

Lee, S.-Y. D., A. M. Arozullah, and Y. I. Cho (2004). Health literacy, social support, and health: a research agenda. *Social Science & Medicine 58*(7), 1309–1321.

Leinonen, R., E. Heikkinen, and M. Jylhä (2002). Changes in health, functional performance and activity predict changes in self-rated health: a 10-year follow-up study in older people. *Archives of Gerontology and Geriatrics 35*(1), 79–92.

Lemire, M., C. Sicotte, and G. Paré (2008). Internet use and the logics of personal empowerment in health. *Health Policy 88*(1), 130–140.

Leventhal, L., I. Brisette, and H. Leventhal (2003). The common-sense model of self-regulation of health and illness. In L. Cameron and H. Leventhal (Eds.), *The self-regulation of health and illness behaviour*. London: Routledge.

Lewis, T. (2006). Seeking health information on the internet: lifestyle choice or bad attack of cyberchondria? *Media, Culture & Society 28*(4), 521–539.

Light, D. W. (2001). Managed competition, governmentality and institutional response in the United Kingdom. *Social Science & Medicine 52*(8), 1167–1181.

Lincoln, Y. S. and E. G. Guba (1985). *Naturalistic Inquiry* (12 ed.). Newbury Park, CA: SAGE Publications.

Lindström, M., B. S. Hanson, and P.-O. Östergren (2001). Socioeconomic differences in leisure-time physical activity: the role of social participation and social capital in shaping health related behaviour. *Social Science & Medicine 52*(3), 441–451.

Lindström, M., B. S. Hanson, P.-O. Östergren, and G. Berglund (2000). Socioeconomic differences in smoking cessation: the role of social participation. *Scandinavian Journal of Public Health 28*(3), 200–208.

Link, B. G. and J. Phelan (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior 35*(Extra Issue: Forty Years of Medical Sociology: The State of the Art and Directions for the Future), 80–94.

Lizardo, O. (2010). Beyond the antinomies of structure: Levi-Strauss, Giddens, Bourdieu, and Sewell. *Theory and Society 39*(6), 651–688.

London Health Observatory (2010). Jubilee line of health inequality 2004-2008. Accessed 04 March 2015, http://www.lho.org.uk/viewResource.aspx?id=15463.

Loos, A. (2013). Cyberchondria: Too Much Information for the Health Anxious Patient? *Journal of Consumer Health On the Internet 17*(4), 439–445.

Lovelace, R. (2014). Introducing spatial microsimulation with R: a practical. Working paper, National Centre for Research Methods (NCRM).

Lovelace, R. and D. Ballas (2013). 'Truncate, replicate, sample': a method for creating integer weights for spatial microsimulation. *Computers 41*(0), 1–11.

Lovelace, R., M. Birkin, D. Ballas, and E. van Leeuwen (2015). Evaluating the performance of iterative proportional fitting for spatial microsimulation: new tests for an established technique. Unpublished work.

Lovelace, R. and M. Dumont (2015). *Spatial Microsimulation with R*. London: Chapman & Hall. forthcoming.

Lowrance, W. (2003). Learning from experience: privacy and the secondary use of data in health research. *Journal of Health Services Research & Policy 8*(suppl 1), 2–7.

Ludwig, E. G. and G. Gibson (1969). Self perception of sickness and the seeking of medical care. *Journal of Health and Social Behavior 10*(2), 125–133.

Lupiáñez, F., I. Maghiros, and F. Abadie (2013). Citizens and ICT for Health in 14 European Countries: Results from an Online Panel. Strategic Intelligence Monitor on Personal Health Systems, Phase 2, European Commission's Joint Research Centre (JRC), Institute for Prospective Technological Studies (IPTS).

Macintyre, S. and A. Ellaway (2003). Neighborhoods and health: an overview. In I. Kawachi and L. F. Berkman (Eds.), *Neighborhoods and Health*, pp. 20–44. Oxford: Oxford University Press.

MacKenzie, D. (1999). The certainty trough. In W. H. Dutton, M. Peltu, and M. Bruce (Eds.), *Society on the Line: Information Politics in the Digital Age*, pp. 549–577. Oxford: Oxford University Press.

MacKenzie, D. A. and J. Wajcman (1999). *The Social Shaping of Technology* (2 ed.). Buckingham: Open University Press.

Mallan, K. M. (2009). Look at me! Look at me! Self-representation and self-exposure through online networks. *Digital Culture and Education 1*(1), 51–56.

Manderson, L., E. Bennett, and S. Andajani-Sutjahjo (2006). The social dynamics of the interview: age, class, and gender. *Qualitative Health Research 16*(10), 1317–1334.

Marmot, M. and R. Wilkinson (2006). *Social determinants of health* (2 ed.). Oxford: Oxford University Press.

Marshall, M. N. (1996). Sampling for qualitative research. *Family Practice 13*(6), 522–526.

Mays, N. and C. Pope (1995). Rigour and qualitative research. *British Medical Journal 311*(6997), 109–112.

McCombs, M. (2005). A look at agenda-setting: past, present and future. *Journalism studies 6*(4), 543–557.

McCombs, M. E. and D. L. Shaw (1972). The agenda-setting function of mass media. *Public Opinion Quarterly 36*(2), 176–187.

McGee, D. L., Y. Liao, G. Cao, and R. S. Cooper (1999). Self-reported health status and mortality in a multiethnic US cohort. *American Journal of Epidemiology 149*(1), 41–46.

McGrenere, J. and W. Ho (2000). Affordances: clarifying and evolving a concept. In *Graphics Interface*, Volume 2000, pp. 179–186.

McKenna, K. and J. A. Bargh (1999). Causes and consequences of social interaction on the Internet: a conceptual framework. *Media Psychology 1*(3), 249–269.

McKenna, K. Y. A. and J. Bargh (1998). Coming out in the age of the Internet: identity 'demarginalization' through virtual group participation. *Journal of Personality and Social Psychology 75*(3), 681–694.

McLellan, E., K. M. MacQueen, and J. L. Neidig (2003). Beyond the qualitative interview: data preparation and transcription. *Field Methods 15*(1), 63–84.

McMullan, M. (2006). Patients using the Internet to obtain health information: how this affects the patient–health professional relationship. *Patient Education and Counseling 63*(1), 24–28.

Mechanic, D. (1972). Social psychologic factors affecting the presentation of bodily complaints. *The New England Journal of Medicine 286*(21), 1132–1139.

Mechanic, D. (1975). Sociocultural and socio-psychological factors affecting personal responses to psychological disorder. *Journal of Health and Social Behavior 16*(4), 393–404.

Mechanic, D. (1978). *Medical Sociology* (2 ed.). New York, NY: Free Press.

Mechanic, D. (1979). Correlates of physician utilization: why do major multivariate studies of physician utilization find trivial psychosocial and organizational effects? *Journal of Health and Social Behavior 20*(4), 387–396.

Merton, R. K. (1957). The role-set: Problems in sociological theory. *British Journal of Sociology 8*(2), 106–120.

Merton, R. K. and P. F. Lazarsfeld (1950). *Continuities in social research: studies in the scope and method of "The American Soldier"*. Glencoe, IL: Free Press.

Meyer, E., R. Schroeder, and U. Deetjen (2016). Dementia data sharing: task certainty and mutual dependence in medical research. *Paper currently under preparation* (forthcoming).

Meyrowitz, J. (1985). *No sense of place: the impact of electronic media on social behavior*. Oxford: Oxford University Press.

Miilunpalo, S., I. Vuori, P. Oja, M. Pasanen, and H. Urponen (1997). Self-rated health status as a health measure: the predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. *Journal of Clinical Epidemiology 50*(5), 517–528.

Miles, M. B. and M. Huberman (1994). *Qualitative Data Analysis: An Expanded Sourcebook* (2 ed.). London; Thousand Oaks, CA: SAGE.

Mirowsky, J. and C. E. Ross (2003). *Education, social status, and health*. New York, NY: Aldine Transaction Publishers.

Mirowsky, J. and C. E. Ross (2008). Education and self-rated health: cumulative advantage and its rising importance. *Research on Aging 30*(1), 93–122.

Mitchell, E. S. (1986). Multiple triangulation: a methodology for nursing science. *Advances in Nursing Science 8*(3), 18–26.

Miyata, K. (2002). Social support for Japanese mothers online and offline. In B. Wellman and C. Haythornthwaite (Eds.), *The Internet in Everyday Life*, Information Age Series. Malden, MA: Blackwell Publishing.

Moran, M. D. (2003). Arguments for rejecting the sequential bonferroni in ecological studies. *Oikos 100*(2), 403–405.

Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: applications to health research. *Qualitative Health Research 8*(3), 362–376.

Morrissey, K., G. Clarke, D. Ballas, S. Hynes, and C. O'Donoghue (2008). Examining access to GP services in rural Ireland using microsimulation analysis. *Area 40*(3), 354–364.

Morrissey, K., G. Clarke, P. Williamson, A. Daly, C. O'Donoghue, et al. (2015). Mental illness in Ireland: simulating its geographical prevalence and the role of access to services. *Environment and Planning B: Planning and Design 42*(2), 338–353.

Morrissey, K., S. Hynes, G. Clarke, and C. O'Donoghue (2010). Examining the factors associated with depression at the small area level in Ireland using spatial microsimulation techniques. *Irish Geography 43*(1), 1–22.

Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing research 40*(2), 120–123.

Morse, J. M. (1995). The significance of saturation. *Qualitative Health Research 5*(2), 147–149.

Mossey, J. M. and E. Shapiro (1982). Self-rated health: a predictor of mortality among the elderly. *American Journal of Public Health 72*(8), 800–808.

MRC (2014). Medical Research Council. Research Changes Lives: MRC Strategic plan 2014–2019. Accessed 17 December 2014, http://www.mrc.ac.uk/news-events/publications/strategic-plan-2014-19/.

Murray, E., B. Lo, L. Pollack, K. Donelan, J. Catania, M. White, K. Zapert, and R. Turner (2003). The impact of health information on the internet on the physician-patient relationship: patient perceptions. *Archives of Internal Medicine 163*(14), 1727–1734.

Mustard, C. A., P. Kaufert, A. Kozyrskyj, and T. Mayer (1998). Sex differences in the use of health care services. *New England Journal of Medicine 338*(23), 1678–1683.

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika 49*(1), 115–132.

Nabi, R. L., A. Prestin, and J. So (2013). Facebook friends with (health) benefits? Exploring social network site use and perceptions of social support, stress, and well-being. *Cyberpsychology, Behavior, and Social Networking 16*(10), 721–727.

Nast, J. and T. Blokland (2014). Social mix revisited: neighbourhood institutions as setting for boundary work and social capital. *Sociology 48*(3), 482–499.

National Library of Medicine (2014). Health Literacy. Technical report.

Nettleton, S., R. Burrows, and L. O'Malley (2005). The mundane realities of the everyday lay use of the internet for health, and their consequences for media convergence. *Sociology of Health & Illness 27*(7), 972–992.

NHS Choices (2014). Hospitals. Accessed 15 May 2014, http://media.nhschoices.nhs.uk/data/foi/Hospital.csv.

NHS Confederation (2015). Key statistics on the NHS. Accessed 17 December 2015, http://www.nhsconfed.org/resources/key-statistics-on-the-nhs.

NHS England (2010). The NHS Atlas of Variation in Healthcare. Accessed 06 March 2015, http://www.rightcare.nhs.uk/atlas/qipp_nhsAtlas-LOW_261110c.pdf.

NHS England (2013). The NHS Constitution: the NHS belongs to us all. Accessed 05 March 2015, http://www.nhs.uk/choiceintheNHS/Rightsandpledges/NHSConstitution/Documents/2013/the-nhs-constitution-for-england-2013.pdf.

NHS England (2014a). High quality datasets: Our vision. Technical report.

NHS England (2014b). Up to 100,000 people use the Internet to improve their health. Accessed 05 March 2015, http://www.england.nhs.uk/2013/03/13/internet-health/.

NIB (2014). National Information Board. Personalised Health and Care 2020: Using Data and Technology to Transform Outcomes for Patients and Citizens. Technical report.

Norman, D. A. (2002). *The design of everyday things*. New York, NY: Basic Books.

Nuffield Council on Bioethics (2015). The Collection, Linking and Use of Data in Biomedical Research and Health Care: Ethical Issues. Technical report.

O'Connor, H., C. Madge, R. Shaw, and J. Wellens (2008). Internet-based interviewing. In N. G. Fielding, R. M. Lee, and G. Blank (Eds.), *The SAGE Handbook of Online Research Methods*, pp. 271–289. London: SAGE.

O'Donoghue, C., N. Farell, K. Morrissey, J. Lennon, D. Ballas, G. Clarke, and S. Hynes (2013). The smile model: construction and calibration. In *Spatial Microsimulation for Rural Policy Analysis*, pp. 55–86. Berlin; Heidelberg: Springer.

Ofcom (2013). Adults' media use and attitudes report 2013. Accessed 23 January 2015, http://stakeholders.ofcom.org.uk/binaries/research/media-literacy/adult-media-lit-13/2013_Adult_ML_Tracker.pdf.

Ofcom (2014). Infrastructure Report 2014: Ofcom's second full analysis of the UK's communications infrastructure. Technical report, London, UK: Ofcom.

Ohm, P. (2010). Broken promises of privacy: responding to the surprising failure of anonymization. *UCLA Law Review 57*(6), 1701–1777.

OLS (2013). Opinions and lifestyle survey, Internet access module, January, February and March. machine-readable dataset of the Office for National Statistics. Accessed 23 January 2015, http://discover.ukdataservice.ac.uk/Catalogue/?sn=7387.

Ong, B. N. (1993). *The Practice of Health Services Research*. London: Chapman & Hall.

ONS (2011a). Office for National Statistics. Household Questionnaire England. Accessed 16 January 2014, http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/2011-census-questionnaire-content/2011-census-questionnaire-for-england.pdf.

ONS (2011b). Office for National Statistics. Statistical disclosure control for 2011 census. Accessed 21 April 2015, http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/processing-the-information/statistical-methodology/statistical-disclosure-control-for-2011-census.pdf.

ONS (2013a). Office for National Statistics. Local authority district (GB) 2013 Boundaries (Generalised, Clipped). Accessed 08 January 2015, https://geoportal.statistics.gov.uk/Docs/Boundaries/Local_authority_district_(GB)_2013_Boundaries_(Generalised_Clipped).zip.

ONS (2013b). Office for National Statistics. Output Area (OA). Accessed 24 November 2013, http://www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/output-area--oas-/index.html.

ONS (2014a). Office for National Statistics. 2011 Census Variable and Classification Information: Part 4 – Derived Variables. Accessed 10 May 2014, http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-user-guide/information-by-variable/part-4--derived-variables.pdf.

ONS (2014b). Office for National Statistics. About the Area Classifications. Accessed 20 June 2014, http://www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/ns-area-classifications/ns-2011-area-classifications/about-the-area-classifications/index.html.

ONS (2014c). Office for National Statistics. The National Statistics Socio-economic Classification: User Manual. Accessed 12 May 2014, http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/section-13--deriving-the-ns-sec--full--reduced-and-simplified-methods.pdf.

Onwuegbuzie, A. J. and K. M. Collins (2007). A typology of mixed methods sampling designs in social science research. *Qualitative Report 12*(2), 281–316.

O'Reilly, M. and N. Parker (2012). "Unsatisfactory Saturation": a critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research*, 1468794112446106.

Orgad, S. (2009). How can researchers make sense of the issues involved in collecting and interpreting online and offline data? In A. Markham and N. Baym (Eds.), *Internet Inquiry: Conversations about Method*, pp. 33–53. Thousand Oaks, CA: SAGE.

OxIS (2003). Oxford Internet Surveys. Machine-readable dataset of the Oxford Internet Institute, University of Oxford.

OxIS (2013). Oxford Internet Surveys. Machine-readable dataset of the Oxford Internet Institute, University of Oxford.

Padfield, M. and I. Procter (1996). The effect of interviewer's gender on the interviewing process: a comparative enquiry. *Sociology 30*(2), 355–366.

Pampalon, R., D. Hamel, M. De Koninck, and M.-J. Disant (2007). Perception of place and health: differences between neighbourhoods in the Québec City region. *Social Science & Medicine 65*(1), 95–111.

Papacharissi, Z. and A. M. Rubin (2000). Predictors of Internet use. *Journal of Broadcasting & Electronic Media 44*(2), 175–196.

Parker, R. (2000). Health literacy: a challenge for American patients and their health care providers. *Health Promotion International 15*(4), 277–283.

Parker, R. M., S. C. Ratzan, and N. Lurie (2003). Health literacy: a policy challenge for advancing high-quality health care. *Health Affairs 22*(4), 147–153.

Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research 4*(5.2), 1189–1208.

Pescosolido, B. A. (1991). Illness careers and network ties: A conceptual model of utilization and compliance. *Advances in Medical Sociology 2*(16), 164–181.

Pescosolido, B. A. (1992). Beyond rational choice: The social dynamics of how people seek help. *American Journal of Sociology 97*(4), 1096–1138.

Pescosolido, B. A., C. B. Gardner, and K. M. Lubell (1998). How people get into mental health services: Stories of choice, coercion and "muddling through" from "first-timers". *Social Science & Medicine 46*(2), 275–286.

Petticrew, M. and H. Roberts (2006). *Systematic Reviews in the Social Sciences*. A Practical Guide. Malden, MA; Oxford: Blackwell.

Pew Research Center (2014). Health fact sheet. Accessed 22 February 2015, http://www.pewinternet.org/fact-sheets/health-fact-sheet/.

Pickett, K. E. and M. Pearl (2001). Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *Journal of Epidemiology and Community Health 55*(2), 111–122.

Pitney, W. and J. Parker (2009). *Qualitative Research in Physical Activity and the Health Professions* (1 ed.). Leeds: Human Kinetics.

Pols, J. and I. Moser (2009). Cold technologies versus warm care? On affective and social relations with and through care technologies. *ALTER, European Journal of Disability Research 3*(2), 159–178.

Pols, J. and D. Willems (2011). Innovation and evaluation: taming and unleashing telecare technology. *Sociology of Health & Illness 33*(3), 484–498.

Poortinga, W. (2006). Perceptions of the environment, physical activity, and obesity. *Social Science & Medicine 63*(11), 2835–2846.

Pope, C. and N. Mays (1995). Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *British Medical Journal 311*, 42–45.

Pope, C. and N. Mays (2006). Qualitative methods in health research. In *Qualitative Research in Health Care* (3 ed.). Malden, MA: Blackwell Publications/BMJ Books.

Pope, C., S. Ziebland, and N. Mays (2006). Analysing qualitative data. In *Qualitative Research in Health Care* (3 ed.). Malden, MA: Blackwell Publications/BMJ Books.

Powell, J., N. Inglis, J. Ronnie, and S. Large (2011). The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. *Journal of Medical Internet Research 13*(1), e20.

Powell, J., P. Lowe, F. Griffiths, and M. Thorogood (2005). A critical analysis of the literature on the Internet and consumer health information. *Journal of Telemedicine and Telecare 11*(suppl 1), 41–43.

Powell, J. A. and S. Boden (2012). Greater choice and control? Health policy in England and the online health consumer. *Policy & Internet 4*(2), 1–23.

Powell, J. A., M. Darvell, and J. A. M. Gray (2003). The doctor, the patient and the world-wide web: how the Internet is changing healthcare. *Journal of the Royal Society of Medicine 96*(2), 74–76.

Power, A. (2000). Poor areas and social exclusion. In A. Power and J. Wilson (Eds.), *Social Exclusion and the Future of Cities*, CASE paper 35, pp. 347–356. London: Centre for Analysis of Social Exclusion.

Primo, D. M., M. L. Jacobsmeier, and J. Milyo (2007). Estimating the impact of state policies and institutions with mixed-level data. *State Politics & Policy Quarterly 7*(4), 446–459.

Prior, L. (2003). Belief, knowledge and expertise: the emergence of the lay expert in medical sociology. *Sociology of Health & Illness 25*(3), 41–57.

Punch, K. F. (1998). *Introduction to Social Research: Quantitative and Qualitative Approaches*. London: SAGE Publications.

Punj, G. and D. W. Stewart (1983). Cluster analysis in marketing research: review and suggestions for application. *Journal of Marketing Research 20*(2), 134–148.

Putnam, R. D. (2000). *Bowling alone: America's declining social capital*. New York, NY; London: Simon & Schuster.

Rabe-Hesketh, S. and A. Skrondal (2012). *Multilevel and Longitudinal Modeling using Stata* (3rd ed.). College Station, TX: Stata Press Publication.

Rahman, A., A. Harding, R. Tanton, and S. Liu (2010). Methodological issues in spatial microsimulation modelling for small area estimation. *International Journal of Microsimulation 3*(2), 3–22.

Rains, S. A. (2007). Perceptions of traditional information sources and use of the world wide web to seek health information: findings from the health information national trends survey. *Journal of Health Communication 12*(7), 667–680.

Rains, S. A. and D. M. Keating (2011). The social dimension of blogging about health: Health blogging, social support, and well-being. *Communication Monographs 78*(4), 511–534.

Rauer, U. (2012). Patient trust in Internet-based health records: an analysis across operator types and levels of patient involvement in germany. *Policy & Internet 4*(2), 1–42.

Rebhan, D. P. (2009). Health care utilization: understanding and applying theories and models of health care seeking behavior. Unpublished work.

Reblin, M. and B. N. Uchino (2008). Social and emotional support and its implication for health. *Current Opinion in Psychiatry 21*(2), 201.

Rees, P., D. Martin, and P. Williamson (2002). Census data resources in the United Kingdom. In *The Census Data System*. Chichester: Wiley.

Reeves, B. and C. I. Nass (1998). *The media equation: how people treat computers, television, and new media like real people and places*. Stanford, CA: CSLI Publications.

Reisdorf, B. C., A.-S. Axelsson, and H. Maurin (2012). Living offline – a qualitative study of Internet non-use in Great Britain and Sweden. *Selected Papers of Internet Research 2*, 1–28.

Reisdorf, B. C. and D. Grošelj (2015). Internet (non-)use types and motivational access: Implications for digital inequalities research. *New Media & Society*, 1461444815621539.

Ribbens, J. (1989). Interviewing – an "unnatural situation"? *Women's Studies International Forum 12*(6), 579–592.

Rice, R. E. (2000). The Internet and health communication: a framework of experiences. In R. E. Rice and J. E. Katz (Eds.), *The Internet and Health Communication: Experiences and Expectations* (1 ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Rice, R. E. (2006). Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. *International Journal of Medical Informatics 75*(1), 8–28.

Rice, R. E. and J. E. Katz (2003). Comparing internet and mobile phone usage: digital divides of usage, adoption, and dropouts. *Telecommunications Policy 27*(8), 597–623.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution 43*(1), 223–225.

Ridd, M., A. Shaw, G. Lewis, and C. Salisbury (2009). The patient–doctor relationship: a synthesis of the qualitative literature on patients' perspectives. *Br J Gen Pract 59*(561), e116–e133.

Rigdon, E. E. (1996). CFI versus RMSEA: a comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal 3*(4), 369–379.

Riva, M., S. Curtis, L. Gauvin, and J. Fagg (2009). Unravelling the extent of inequalities in health across urban and rural areas: evidence from a national sample in England. *Social Science & Medicine 68*(4), 654–663.

Roberts, K. J. (1999). Patient empowerment in the United States: a critical commentary. *Health Expectations 2*(2), 82–92.

Robinson, L., S. R. Cotten, H. Ono, A. Quan-Haase, G. Mesch, W. Chen, J. Schulz, T. M. Hale, and M. J. Stern (2015). Digital inequalities and why they matter. *Information, Communication & Society 18*(5), 569–582.

Rogers, A. and N. Mead (2004). More than technology and access: primary care patients' views on the use and non-use of health information in the Internet age. *Health & Social Care in the Community 12*(2), 102–110.

Rogers, A. and G. Nicolaas (1998). Understanding the patterns and processes of primary care use: a combined quantitative and qualitative approach. *Sociological Research Online 3*(4).

Rose, R. (2000). How much does social capital add to individual health? *Social Science & Medicine 51*(9), 1421–1435.

Rosenstock, I. M. (1966). Why people use health services. *The Milbank Memorial Fund Quarterly 44*(3), 94–127. Part 2: Health Services Research I.

Rosenstock, I. M. (1974). Historical origins of the health belief model. *Health Education & Behavior 2*(4), 328–335.

Rosenstock, I. M., V. J. Strecher, and M. H. Becker (1988). Social learning theory and the health belief model. *Health Education & Behavior 15*(2), 175–183.

Rubin, A. M. (2009). Uses-and-gratifications perspective on media effects. In J. Bryant and M. B. Oliver (Eds.), *Media Effects: Advances in Theory and Research* (3 ed.). New York, NY; London: Routledge.

Rubin, H. J. and I. S. Rubin (2004). *Qualitative Interviewing: The Art of Hearing Data* (2 ed.). Los Angeles, CA; London: SAGE.

Sadasivam, R. S., R. L. Kinney, S. C. Lemon, S. L. Shimada, J. J. Allison, and T. K. Houston (2013). Internet health information seeking is a team sport: analysis of the pew internet survey. *International Journal of Medical Informatics 82*(3), 193–200.

Saldaña, J. (2012). *The Coding Manual for Qualitative Researchers* (2 ed.). London: SAGE.

Samal, L., S. Saha, and G. Chander (2011). Internet health information seeking behavior and antiretroviral adherence in persons living with HIV/AIDS. *AIDS Patient Care and STDs 25*(7), 445–449.

Sandelowski, M. (1995a). Qualitative analysis: what it is and how to begin. *Research in Nursing & Health 18*(4), 371–375.

Sandelowski, M. (1995b). Sample size in qualitative research. *Research in Nursing & Health 18*(2), 179–183.

Santana, S., B. Lausen, M. Bujnowska-Fedak, C. E. Chronaki, H.-U. Prokosch, and R. Wynn (2011). Informed citizen and empowered citizen in health: results from an European survey. *BMC Family Practice 12*(1), 20.

Savage, M. (2013). Digital fields, networks and capital: Sociology beyond structures and fluids. In *Digital Sociology: Critical Perspectives*, pp. 139–150. Basingstoke: Palgrave Macmillan.

Savolainen, R. (1995). Everyday life information seeking: approaching information seeking in the context of "way of life". *Library & Information Science Research 17*(3), 259–294.

Schaefer, C., J. C. Coyne, and R. S. Lazarus (1981). The health-related functions of social support. *Journal of Behavioral Medicine 4*(4), 381–406.

Schaffer, R., K. Kuczynski, and D. Skinner (2008). Producing genetic knowledge and citizenship through the Internet: mothers, pediatric genetics, and cybermedicine. *Sociology of Health & Illness 30*(1), 145–159.

Schneider, G., G. Driesch, A. Kruse, M. Wachter, H.-G. Nehen, and G. Heuft (2004). What influences self-perception of health in the elderly? The role of objective health condition, subjective well-being and sense of coherence. *Archives of Gerontology and Geriatrics 39*(3), 227–237.

Seale, C. (1999). Quality in qualitative research. *Qualitative Inquiry 5*(4), 465–478.

Seidman, I. (2013). *Interviewing as qualitative research: a guide for researchers in education and the social sciences* (4 ed.). New York, NY: Teachers College Press.

Shalizi, C. (2009). Distances between Clustering, Hierarchical Clustering. Accessed 15 January 2015, http://www.stat.cmu.edu/~cshalizi/350/lectures/08/lecture-08.pdf.

Sillence, E., P. Briggs, P. R. Harris, and L. Fishwick (2007). How do patients evaluate and make use of online health information? *Social Science & Medicine 64*(9), 1853–1862.

Silverman, D. (2006). *Interpreting Qualitative Data: Methods for Analyzing Talk, Text and Interaction* (3rd ed.). London: SAGE.

Silverman, D. (2013). *Doing Qualitative Research: A Practical Handbook*. London: SAGE Publications Limited.

Singh, K. (2007). *Quantitative Social Research Methods*. Los Angeles, CA; London: SAGE.

Smith, D. M., G. P. Clarke, and K. Harland (2009). Improving the synthetic data generation process in spatial microsimulation models. *Environment and Planning A 41*(5), 1251.

Smith, D. M., J. R. Pearce, and K. Harland (2011). Can a deterministic spatial microsimulation model provide reliable small-area estimates of health behaviours? An example of smoking prevalence in New Zealand. *Health & Place 17*(2), 618–624.

Song, M., M. Sandelowski, and M. Happ (2010). Current practices and emerging trends in conducting mixed methods intervention studies in the health sciences. In *SAGE Handbook of Mixed Methods in Social & Behavioral Research*, pp. 725–748. Thousand Oaks, CA: SAGE.

Soobader, M.-J. and F. B. LeClere (1999). Aggregation and the measurement of income inequality: effects on morbidity. *Social Science & Medicine 48*(6), 733–744.

Spencer, L., J. Ritchie, W. O'Connor, G. Morell, and R. Ormston (2013). Analysis in practice. In J. Ritchie, J. Lewis, C. M. Nicholls, and R. Ormston (Eds.), *Qualitative research practice: A guide for social science students and researchers* (2 ed.)., pp. 295–346. London: SAGE Publications.

Spielman, S. E. and E.-h. Yoo (2009). The spatial dimensions of neighborhood effects. *Social Science & Medicine 68*(6), 1098–1105.

Stafford, T. F., M. R. Stafford, and L. L. Schkade (2004). Determining uses and gratifications for the Internet. *Decision Sciences 35*(2), 259–288.

Stevens, J., J. Cai, K. R. Evenson, and R. Thomas (2002). Fitness and fatness as predictors of mortality from all causes and from cardiovascular disease in men and women in the lipid research clinics study. *American Journal of Epidemiology 156*(9), 832–841.

Stevens, L. A. and G. Laurie (2014). The administrative data research centre scotland: a scoping report on the legal & ethical issues arising from access & linkage of administrative data. *Edinburgh School of Law Research Paper* (2014/35).

Stockdale, S. E., K. B. Wells, L. Tang, T. R. Belin, L. Zhang, and C. D. Sherbourne (2007). The importance of social context: neighborhood stressors, stress-buffering mechanisms, and alcohol, drug, and mental health disorders. *Social Science & Medicine 65*(9), 1867–1881.

Straubhaar, J., Z. Tufekci, J. Spence, and V. Rojas (2012). Digital inequity in the Austin Technopolis: An introduction. In J. Straubhaar, J. Spence, Z. Tufekci, and R. G. Lentz (Eds.), *Inequity in the Technopolis: race, class, gender, and the digital divide in Austin*. Austin, TX: University of Texas Press.

Suchman, E. A. (1965). Stages of illness and medical care. *Journal of Health and Human Behavior 6*(3), 114–128.

Sullivan, G. M. and R. Feinn (2012). Using effect size – or why the p-value is not enough. *Journal of Graduate Medical Education 4*(3), 279–282.

Suominen, S., H. Helenius, H. Blomberg, A. Uutela, and M. Koskenvuo (2001). Sense of coherence as a predictor of subjective state of health: results of 4 years of follow-up of adults. *Journal of Psychosomatic Research 50*(2), 77–86.

Suri, H. (2011). Purposeful sampling in qualitative research synthesis. *Qualitative Research Journal 11*(2), 63–75.

Sutherland, L. R. and M. J. Verhoef (1994). Why do patients seek a second opinion or alternative medicine? *Journal of Clinical Gastroenterology 19*(3), 194–197.

Suziedelyte, A. (2012). How does searching for health information on the Internet affect individuals' demand for health care services? *Social Science & Medicine 75*(10), 1828–1835.

Tang, E. and W. Lee (2006). Singapore Internet users' health information search: motivation, perception of information sources, and self-efficacy. In M. Murero and R. E. Rice (Eds.), *The Internet and Health Care: Theory, Research, and Practice*, pp. 107–126. Mahwah, NJ: Lawrence Erlbaum Associates.

Tanton, R. and K. L. Edwards (2013). Introduction to spatial microsimulation: history, methods and applications. In R. Tanton and K. L. Edwards (Eds.), *Spatial Microsimulation: A Reference Guide for Users*, pp. 3–8. Dordrecht: Springer.

Tashakkori, A. and C. Teddlie (1998). *Mixed methodology: combining qualitative and quantitative approaches*. Thousand Oaks, CA; London: SAGE.

Taylor, D. and M. Bury (2007). Chronic illness, expert patients and care transition. *Sociology of Health & Illness 29*(1), 27–45.

Teddlie, C. and F. Yu (2007). Mixed methods sampling: a typology with examples. *Journal of Mixed Methods Research 1*(1), 77–100.

The King's Fund (2014). The autumn statement: NHS Funding. Accessed 05 March 2015, http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/autumn-statement-briefing-2014.pdf.

Thoits, P. A. (2013). Self, identity, stress, and mental health. In C. S. Aneshensel, J. C. Phelan, and A. Bierman (Eds.), *Handbook of the sociology of mental health*, pp. 357–377. Dordrecht; New York, NY: Springer.

Thompson, J. B. (1995). *The media and modernity: A social theory of the media*. Stanford, CA: Stanford University Press.

Tian, Y. and J. D. Robinson (2009). Incidental health information use on the Internet. *Health Communication 24*(1), 41–49.

Tiggemann, M. and B. McGill (2004). The role of social comparison in the effect of magazine advertisements on women's mood and body dissatisfaction. *Journal of Social and Clinical Psychology 23*(1), 23–44.

Tinder Foundation (2013). Tinder Foundation wins NHS contract. Accessed 29 November 2013, http://www.tinderfoundation.org/our-thinking/news/tinder-foundation-wins-nhs-contract.

Tinder Foundation (2014). Improving access to digital health resources: A report on the Widening Digital Participation project. Technical report, Presented by Tinder Foundation, in association with NHS England.

Tjora, A. (2014). Digital patients: A typology of emerging patient roles. In E. A. A. Jaatun, E. Brooks, K. E. Berntsen, H. Gilstad, and M. G. Jaatun (Eds.), *Proceedings of the 2nd European Workshop on Practical Aspects of Health Informatics (PAHI 2014), Trondheim, Norway*, pp. 135–144.

Tomintz, M. N., G. P. Clarke, and J. E. Rigby (2008). The geography of smoking in Leeds: estimating individual smoking rates and the implications for the location of stop smoking services. *Area 40*(3), 341–353.

Tracy, S. J. (2010). Qualitative quality: Eight big-tent criteria for excellent qualitative research. *Qualitative Inquiry 16*(10), 837–851.

Treiber, F. A., T. Baranowski, D. S. Braden, W. B. Strong, M. Levy, and W. Knox (1991). Social support for exercise: Relationship to physical activity in young adults. *Preventive Medicine 20*(6), 737–750.

Tucker, K. H. (1998). *Anthony Giddens and modern social theory*. Thousand Oaks, CA: SAGE.

Tunstall, H. V. Z., M. Shaw, and D. Dorling (2004). Places and health. *Journal of Epidemiology and Community Health 58*(1), 6–10.

Turkle, S. (1997). Multiple subjectivity and virtual community at the end of the Freudian century. *Sociological Inquiry 67*(1), 72–84.

Turner, J. H. (2005). A new approach to theoretically integrating micro and macro analysis. In C. Calhoun, C. Rojek, and B. S. Turner (Eds.), *The SAGE Handbook of Sociology*, pp. 405–422. London; Thousand Oaks, CA: SAGE.

Tustin, N. (2010). The role of patient satisfaction in online health information seeking. *Journal of Health Communication 15*(1), 3–17.

Uchino, B. N. (2004). *Social support and physical health: Understanding the health consequences of relationships*. New Haven, CT: Yale University Press.

Uchino, B. N. (2006). Social support and health: a review of physiological processes potentially underlying links to disease outcomes. *Journal of Behavioral Medicine 29*(4), 377–387.

Uchino, B. N., J. T. Cacioppo, and J. K. Kiecolt-Glaser (1996). The relationship between social support and physiological processes: a review with emphasis on underlying mechanisms and implications for health. *Psychological Bulletin 119*(3), 488.

UK Government (2002). Too soon to scrap the census. Technical report, Fifteenth Report of Session 201314.

UK Government (2013). Government Digital Strategy: December 2013. Technical report, Cabinet Office.

UK Government (2014). Government Digital Inclusion Strategy. Technical report, Cabinet Office & Government Digital Services.

UK National Audit Office (2013). Digital Britain 2: Putting users at the heart of government's digital services. Accessed 11 February 2014, http://www.nao.org.uk/wp-content/uploads/2013/07/10123-001-Digital-Britain-2-Book.pdf.

UKHLS (2013). UK Household Longitudinal Study. Machine-readable dataset of the University of Essex. Accessed 27 January 2015, http://discover.ukdataservice.ac.uk/catalogue/?sn=6614.

van Deursen, A. and J. van Dijk (2011a). Internet skills and the digital divide. *New Media & Society 13*(6), 893–911.

van Deursen, A. J. and E. J. Helsper (2015). A nuanced understanding of internet use and non-use among the elderly. *European Journal of Communication 30*(2), 171–187.

van Deursen, A. J. and J. A. van Dijk (2011b). Internet skills performance tests: are people ready for ehealth? *Journal of Medical Internet Research 13*(2), e35.

van Deursen, A. J. and J. A. van Dijk (2014). The digital divide shifts to differences in usage. *New Media & Society 16*(3), 507–526.

van Deursen, A. J., J. A. van Dijk, and O. Peters (2011). Rethinking Internet skills: the contribution of gender, age, education, Internet experience, and hours online to medium-and content-related Internet skills. *Poetics 39*(2), 125–144.

van Deursen, A. J. A. M. and E. J. Helsper (2015). The third-level digital divide: who benefits most from being online? In L. Robinson, S. R. Cotten, J. Schulz, T. M. Hale, and A. Williams (Eds.), *Communication and Information Technologies Annual: Digital Distinctions and Inequalities*, Volume 10, pp. 29–52. Emerald Group Publishing Limited.

van Dijk, J. and K. Hacker (2003). The digital divide as a complex and dynamic phenomenon. *The Information Society 19*(4), 315–326.

van Dijk, J. A. (1999). *The network society: social aspects of new media*. London: SAGE.

van Dijk, J. A. (2005). *The deepening divide: Inequality in the information society*. Thousand Oaks, CA; London: SAGE Publications.

Voas, D. and P. Williamson (2001). Evaluating goodness-of-fit measures for synthetic microdata. *Geographical and Environmental Modelling 5*(2), 177–200.

Wald, H. S., C. E. Dube, and D. C. Anthony (2007). Untangling the web – the impact of Internet use on health care and the physician–patient relationship. *Patient Education and Counseling 68*(3), 218–224.

Wallsten, K. (2007). Agenda setting and the blogosphere: An analysis of the relationship between mainstream media and political blogs. *Review of Policy Research 24*(6), 567–587.

Walther, J. B. (1996). Computer-mediated Communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research 23*(1), 3–43.

Walther, J. B. and S. Boyd (2002). Attraction to computer-mediated social support. In C. A. Lin and D. J. Atkin (Eds.), *Communication Technology and Society: Audience Adoption and Uses*, The Hampton Press communication series: New media, policy and social research issues, pp. 153–188. Cresskill, NJ: Hampton Press.

Wan, T. T. (1976). Predicting self-assessed health status: a multivariate approach. *Health Services Research 11*(4), 464.

Wang, Y. C., R. E. Kraut, and J. M. Levine (2015). Eliciting and receiving online support: Using computer-aided content analysis to examine the dynamics of online social support. *Journal of Medical Internet Research 17*(4), e99.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association 58*(301), 236–244.

Warschauer, M. (2003). Dissecting the "digital divide": A case study in Egypt. *The Information Society 19*(4), 297–304.

Warschauer, M. (2004). *Technology and Social Inclusion: Rethinking the Digital Divide*. Cambridge, MA: MIT press.

Wasserstein, R. L. and N. A. Lazar (2016). The ASA's statement on p-values: context, process, and purpose. *The American Statistician 70*(2), 129–133.

Watt, H. (2015). Jeremy Hunt: People will be diagnosed at home by a computer. Accessed 05 February 2015, http://www.telegraph.co.uk/news/nhs/11391132/Jeremy-Hunt-People-will-be-diagnosed-at-home-by-a-computer.html.

Weber, D. O. and C. P. O. Entry (2004). Survey reveals physicians' love/hate relationship with technology. *The Physician Executive 30*(2), 4–10.

Weiss, R. S. (1995). *Learning from Strangers: The Art and Method of Qualitative Interview Studies*. New York, NY: The Free Press.

Wellman, B. and K. Frank (2001). Network capital in a multi-level world: Getting support from personal communities. In N. Lin, K. S. Cook, and R. S. Burt (Eds.), *Social capital: theory and research* (4 ed.)., pp. 233–273. Chicago, IL: Aldine De Gruyter.

Wellman, B., A. Q. Haase, J. Witte, and K. Hampton (2001). Does the Internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment. *American Behavioral Scientist 45*(3), 436–455.

Wellman, B. and K. Hampton (1999). Living networked on and offline. *Contemporary Sociology 28*(6), 648–654.

Wellman, B., A. Quan-Haase, J. Boase, W. Chen, K. Hampton, I. Díaz, and K. Miyata (2003). The social affordances of the Internet for networked individualism. *Journal of Computer-Mediated Communication 8*(3), 0–0.

Wethington, E. and R. C. Kessler (1986). Perceived support, received support, and adjustment to stressful life events. *Journal of Health and Social behavior 27*(1), 78–89.

Whetten, K., J. Leserman, R. Whetten, J. Ostermann, N. Thielman, M. Swartz, and D. Stangl (2006). Exploring lack of trust in care providers and the government as a barrier to health service use. *American Journal of Public Health 96*(4), 716.

Whitacre, B. and L. Brooks (2014). Do broadband adoption rates impact a community's health? *Behaviour & Information Technology 33*(7), 767–779.

White, R. W. and E. Horvitz (2009). Cyberchondria: studies of the escalation of medical concerns in web search. *ACM Transactions on Information Systems (TOIS) 27*(4), 23:1–23:37.

Whitworth, A. (2013). Evaluations and Improvements in Small Area Estimation Methodologies. Methodological Review paper, National Centre for Research Methods (NCRM). Accessed 07 February 2014, http://eprints.ncrm.ac.uk/3210/1/sme_whitworth.pdf.

WHO (1946). *Constitution of the World Health Organization* (45 ed.). Basic Documents, Supplement, October 2006. Accessed 13 May 2015, http://www.who.int/governance/eb/who_constitution_en.pdf.

Wilkinson, L., G. Blank, and C. Gruber (1996). *Desktop Data Analysis SYSTAT*. Upper Saddle River, NJ; London: Prentice Hall.

Williamson, P. (2013). Combinatorial Optimisation. In A. Whitworth (Ed.), *Evaluations and Improvements in Small Area Estimation Methodologies*, Methodological Review paper. Accessed 07 February 2014, http://eprints.ncrm.ac.uk/3210/1/sme_whitworth.pdf: National Centre for Research Methods (NCRM).

Wilson, I. B. and P. D. Cleary (1995). Linking clinical variables with health-related quality of life: a conceptual model of patient outcomes. *Journal of the American Medical Association 273*(1), 59–65.

Wilson, P. M., S. Kendall, and F. Brooks (2007). The expert patients programme: a paradox of patient empowerment and medical dominance. *Health & Social Care in the Community 15*(5), 426–438.

Winship, C. and S. L. Morgan (1999). The estimation of causal effects from observational data. *Annual Review of Sociology 25*, 659–706.

Wong, D. W. (1992). The reliability of using the iterative proportional fitting procedure. *The Professional Geographer 44*(3), 340–348.

Wray, N., M. Markovic, and L. Manderson (2007). "Researcher Saturation": The impact of data triangulation and intensive-research practices on the researcher and qualitative research process. *Qualitative Health Research 17*(10), 1392–1402.

Wright, K. B. and S. B. Bell (2003). Health-related support groups on the Internet: Linking empirical findings to social support and computer-mediated communication theory. *Journal of Health Psychology 8*(1), 39–54.

Wyatt, S., F. Henwood, A. Hart, and J. Smith (2005). The digital divide, health information and everyday life. *New Media & Society 7*(2), 199–218.

Wyke, S., J. Adamson, D. Dixon, and K. Hunt (2013). Consultation and illness behaviour in response to symptoms: a comparison of models from different disciplinary frameworks and suggestions for future research directions. *Social Science & Medicine 86*, 79–87.

Xiao, N., R. Sharman, H. R. Rao, and S. Upadhyaya (2014). Factors influencing online health information search: An empirical analysis of a national cancer-related survey. *Decision Support Systems 57*, 417–427.

Yates, F. (1934). Contingency tables involving small numbers and the χ 2 test. *Supplement to the Journal of the Royal Statistical Society 1*(2), 217–235.

Ybarra, M. and M. Suman (2008). Reasons, assessments and actions taken: sex and age differences in uses of Internet health information. *Health Education Research 23*(3), 512–521.

Ybarra, M. L. and M. Suman (2006). Help seeking behavior and the Internet: a national survey. *International Journal of Medical Informatics 75*(1), 29–41.

Yeo, A., R. Legard, J. Keegan, K. Ward, C. McNaughton Nicholls, and J. Lewis (2013). In-depth interviews. In J. Ritchie, J. Lewis, C. M. Nicholls, and R. Ormston (Eds.), *Qualitative research practice: A guide for social science students and researchers* (2 ed.)., pp. 177–210. London: SAGE Publications.

Young, J. C. (1981). *Medical choice in a Mexican Village*. New Brunswick, NJ: Rutgers University Press.

Yu, L. (2006). Understanding information inequality: Making sense of the literature of the information and digital divides. *Journal of Librarianship and Information Science 38*(4), 229–252.

Zhao, Y. and G. Karypis (2004). Empirical and theoretical comparisons of selected criterion functions for document clustering. *Machine Learning 55*(3), 311–331.

Ziebland, S. (2004). The importance of being expert: the quest for cancer information on the Internet. *Social Science & Medicine 59*(9), 1783–1793.

Ziebland, S., A. Chapple, C. Dumelow, J. Evans, S. Prinjha, and L. Rozmovits (2004). How the Internet affects patients' experience of cancer: a qualitative study. *British Medical Journal 328*(564).

Ziebland, S. and S. Wyke (2012). Health and illness in a connected world: how might sharing experiences on the Internet affect people's health? *Milbank Quarterly 90*(2), 219–249.

Zubin, J. (1938). A technique for measuring like-mindedness. *The Journal of Abnormal and Social Psychology 33*(4), 508.