Table of contents

List of Tables ........................................................................................................................................... 7
List of Figures ........................................................................................................................................... 10
Glossary and acronyms ............................................................................................................................... 13
Abstract ................................................................................................................................................... 14
Lay Abstract ............................................................................................................................................ 15
Declaration ............................................................................................................................................. 16
Copyright Statement ................................................................................................................................. 16
Acknowledgements ................................................................................................................................. 18
The Author ............................................................................................................................................... 19

Chapter 1: Introduction ............................................................................................................................... 20

1.1 Overview ............................................................................................................................................. 20
1.2 Why focus on patient factors in lung cancer? ...................................................................................... 21
1.3 Help-seeking behaviour ....................................................................................................................... 24
  1.3.1 What is ‘help-seeking behaviour’? .................................................................................................. 24
  1.3.2 Measuring help-seeking .............................................................................................................. 28
  1.3.3 Help-seeking and cancer .............................................................................................................. 29
  1.3.4 Previous interventions for help-seeking behaviour in cancer ..................................................... 30
1.4 The Internet, the World Wide Web and Health ..................................................................................... 31
1.5 In summary ......................................................................................................................................... 33
1.6 Thesis aim .......................................................................................................................................... 33
1.7 Thesis structure ................................................................................................................................. 34

Chapter 2: Web use for symptom appraisal for physical health conditions: a systematic review ................. 35

2.1 Introduction ......................................................................................................................................... 35
2.2 Methods ............................................................................................................................................. 36
  2.2.1 Eligibility criteria .......................................................................................................................... 37
  2.2.2 Information sources ..................................................................................................................... 38
  2.2.3 Search .......................................................................................................................................... 38
  2.2.4 Study selection ............................................................................................................................. 38
  2.2.5 Data collection process and data items ....................................................................................... 39
  2.2.6 Quality appraisal .......................................................................................................................... 39
  2.2.7 Synthesis of results ..................................................................................................................... 39
2.3 Results ................................................................................................................................................. 40
  2.3.1 Study selection ............................................................................................................................. 40
  2.3.2 Study characteristics .................................................................................................................... 41
  2.3.3 Quality assessment and risk of bias ............................................................................................ 42
  2.3.4 RQ1: What proportion of different populations use the Web to appraise symptoms? ......... 59
  2.3.5 RQ2: Which symptoms are likely to be researched online? ..................................................... 60
  2.3.6 RQ3: How is Web use for symptom appraisal conducted (search strategies)? .................... 61
  2.3.7 RQ4: Behavioural consequences of Web use for symptom appraisal ...................................... 66
  2.3.8 RQ5: Emotional consequences of Web use for symptom appraisal ....................................... 69
2.4 Discussion .......................................................................................................................................... 70
  2.4.1 Web use for symptom appraisal – assistance or hindrance to health promotion? ........ 71
  2.4.2 Recommendations to improve online health information ....................................................... 73
Chapter 3: The role of Web-based health information in help-seeking behaviour prior to a diagnosis of lung cancer: A mixed-methods study .............................................................. 78

3.1 Introduction ................................................................................. 78
3.1.1 The pathway to diagnosis .................................................... 78
3.1.2 Aims and research questions ................................................ 80
3.2 Method ...................................................................................... 80
3.2.1 Design .................................................................................. 80
3.2.2 Participants and recruitment ................................................ 81
3.2.3 Procedure ............................................................................ 81
3.2.4 Measures ............................................................................ 82
3.2.5 Analyses ................................................................................ 83
3.3 Survey results ............................................................................ 85
3.3.1 Sample description .............................................................. 85
3.3.2 Use of lay information sources ............................................ 89
3.3.3 Proportion of people with lung cancer who report Web searches prior to diagnosis 90
3.3.4 Description of Web searches ................................................ 91
3.4 Qualitative interview results ...................................................... 92
3.4.1 Perceptions of the impact of pre-diagnosis Web searches on the pathway to diagnosis .................................................. 93
3.4.2 Perceived barriers to using the Web pre-diagnosis ............... 95
3.4.3 Search strategies and information appraisal .......................... 96
3.4.4 Use of other lay information sources .................................... 97
3.5 Discussion .................................................................................. 99
3.5.1 Proportion of people with lung cancer reporting pre-diagnosis Web searches .... 99
3.5.2 Perceived impacts of Web searches on the pathway to diagnosis ............. 100
3.5.3 Barriers to pre-diagnosis Web searches: Current and future importance of the Web prior to a lung cancer diagnosis .......... 102
3.5.4 Recommendations and future research ............................... 103
3.5.5 Limitations .......................................................................... 104
3.6 Conclusions ............................................................................. 104

Chapter 4: Development of a Web-based intervention to encourage early help-seeking in people with symptoms associated with lung cancer .................................................. 106

4.1 Introduction ............................................................................. 106
4.2 Overview of the intervention development process .................. 107
4.3 Preliminary work and preparations ....................................... 109
4.3.1 Drawing on qualitative research with users and previous literature ...................... 109
4.3.2 Selecting a theoretical approach ........................................ 111
4.3.3 Combining findings from interviews with theory .......................... 115
4.3.4 Patient and public involvement .......................................... 115
4.3.5 Guiding principles ............................................................... 116
4.4 Incorporating TPB into the intervention ..................................... 118
4.4.1 Identification of relevant beliefs and linking these to TPB constructs .................. 118
4.4.2 Identification of behaviour change techniques ....................... 120
4.4.3 Operationalising behaviour change techniques ...................... 121
4.5 Tailoring ................................................................................... 126
4.5.1 Symptoms and risk factors ................................................ 126
4.5.2 TPB constructs ................................................................. 128
4.5.3 Affected individual or proxy? .............................................. 129
4.6 Medical content ....................................................................... 129
4.7 Think Aloud evaluation ............................................................ 131
4.7.1 Participants ......................................................................... 131
4.7.2 Procedure ........................................................................... 131
Chapter 7: Trial 2: Testing a tailored, theory-based intervention to encourage help-seeking for lung cancer symptoms: A two-way mixed factorial design ........................................ 196

7.1 Aim .................................................................................................................. 196
7.2 Methods .......................................................................................................... 196
7.2.1 Intervention changes .................................................................................. 196
7.2.2 Study design ............................................................................................... 196
7.2.3 Participants .................................................................................................. 197
7.2.4 Outcome measures ...................................................................................... 197
7.2.5 Procedure .................................................................................................... 200
7.2.6 Data handling and coding .......................................................................... 200
7.2.7 Hypotheses .................................................................................................. 202
7.2.8 Statistical analyses ...................................................................................... 202
7.3 Think Aloud evaluation .................................................................................. 204
7.3.1 Participants .................................................................................................. 204
7.3.2 Procedure .................................................................................................... 204
7.3.3 Analysis ....................................................................................................... 204
7.3.4 Results ......................................................................................................... 204
7.3.5 Conclusions for the intervention/trial design ............................................. 206
7.4 Results ............................................................................................................. 206
7.4.1 Sample description ..................................................................................... 206
7.4.2 Primary outcome (self-reported likelihood of visiting a doctor) ................. 210
7.4.3 Follow-up .................................................................................................... 215
7.4.4 Secondary outcomes .................................................................................. 216
7.4.5 Group comparison on sociodemographic variables ..................................... 218
7.5 Discussion ....................................................................................................... 219
7.5.1 Main findings .............................................................................................. 219
7.5.2 Limitations .................................................................................................. 223
7.6 Conclusions ................................................................................................... 225

Chapter 8: Overall discussion ............................................................................... 226

8.1 Overview ......................................................................................................... 226
8.2 Main findings from this thesis ........................................................................ 228
8.3 Strengths and limitations of the studies presented within this thesis ............. 230
8.4 Implications ..................................................................................................... 234
8.5 Future research ............................................................................................... 237
8.6 Conclusion ....................................................................................................... 239

References ........................................................................................................... 242
Appendix 1: Database search strategy ................................................................. 262
Appendix 2: Data extraction sheet ....................................................................... 263
Appendix 3: Quality appraisal of studies included in the systematic review ......... 265
Appendix 4: NVivo matrix example for the systematic review ......................... 287
Appendix 5: NHS Ethics approval letter .............................................................. 288
Appendix 6: Participant Information Sheet for mixed-methods study ................. 289
Appendix 7: Questionnaire/survey ...................................................................... 294
Appendix 8: Interview protocol .......................................................................... 305
Appendix 9: NVivo matrix example for Framework Analysis ......................... 311
Appendix 10: UREC approval letter ................................................................. 312
Appendix 11: Participant Information Sheet (Feasibility study) ....................... 313
Appendix 12: Initial survey (Feasibility study) .................................................. 317
Appendix 13: TPB-components item ................................................................. 320
Appendix 14: Button-click measure (Feasibility study) ..................................... 321
Appendix 15: Intention measure (Feasibility study) ........................................... 322
Appendix 16: Demographic measures ............................................................... 323
Appendix 17: TPB questionnaire ........................................................................ 324
Appendix 18: Risk perception scale ................................................................... 327
Appendix 19: When2Go evaluation .................................................................... 328
Appendix 20: Initial survey (Trial 1) .......................................................................................................................... 329
Appendix 21: Intention measure (Trial 1) ..................................................................................................................... 333
Appendix 22: Next steps measure (Trial 1) ................................................................................................................... 334
Appendix 23: Participant Information Sheet (PIS) (Trial 1) .......................................................................................... 335
Appendix 24: Re-run of non-parametric ANOVAs with randomly selected groups .................................................. 339
Appendix 25: Self-reported likelihood of seeing a doctor about symptoms (Trial 2) ................................................. 348
Appendix 26: Initial survey (Trial 2) ............................................................................................................................ 349
Appendix 27: Next steps measure (Trial 2) .................................................................................................................. 351
Appendix 28: TPB questionnaire (Trial 2) .................................................................................................................... 352
Appendix 29: Follow-up questionnaire (Trial 2) ........................................................................................................... 355
Appendix 30: Participant Information Sheet (PIS) (Trial 2) ......................................................................................... 359

Word count (excluding tables, figures and references): 52,840
List of Tables

Table 2.1 Study characteristics of the studies included in the review. ................................................. 43
Table 2.2. Symptoms and diagnosis examined in included studies. ....................................................... 55
Table 2.3. Percentage of people engaging in Web use for symptom appraisal reported by included studies. ............................................................................................................................................................. 60
Table 3.1 Participant demographic details. ................................................................................................. 87
Table 3.2 Habitual Web/technology use among participants. ........................................................................ 88
Table 3.3 Search terms used during Web searches prior to diagnosis. ................................................... 92
Table 4.1 Incorporation of findings from the interview study and systematic review into the development of the intervention. ................................................................................................................................. 110
Table 4.2 Barriers to help-seeking in lung cancer identified in previous literature. ......................... 112
Table 4.3 Beliefs about help-seeking identified from the literature, mapped onto TPB constructs. ........................................................................................................................................................................... 120
Table 4.4 Construct domains and relevant behaviour change techniques identified from (Michie et al., 2008). ................................................................................................................................................. 122
Table 4.5 Behaviour change techniques (BCTs) identified as suitable and likely to be effective in the TPB-based intervention ........................................................................................................... 124
Table 4.6 Statements used to tailor TPB components to individual users. ........................................... 128
Table 4.7 Details of Web pages used to inform the intervention content. ............................................. 129
Table 4.8 Summary of feedback received and observations made during the Think Aloud evaluation. ........................................................................................................................................................................... 133
Table 5.1 Description of the four study groups. ...................................................................................... 139
Table 5.2 Statements used to tailor TPB components to individual users. ........................................... 143
Table 5.3 Design and structure of the TPB questionnaire. ......................................................................... 146
Table 5.4 Overview of data collected by Google Analytics service. .................................................... 149
Table 5.5 Self-reported demographic data of the sample (N = 130). .................................................. 152
Table 5.6 Demographic data of those experiencing the symptoms (either self-reported or reported by proxy). ................................................................................................................................................. 153
Table 5.7 Number and percentage of participants selecting each of the four TPB-statements. ........................................................................................................................................................................ 154
Table 5.8 Number of clicks on the study link via the Google Ad Words campaign, grouped by search terms entered by users into Google. ................................................................. 155
Table 5.9 Comparison between the four study groups on categorical demographic variables using chi-square test. .................................................................................................. 156
Table 5.10 Number of people who clicked the button-click measure appropriately. .......... 156
Table 5.11 Number of people who responded to the intention item appropriately. .......... 157
Table 5.12 Comparison between the four study groups on TPB constructs. ....................... 158
Table 5.13 Mean intention (7-point scale) by study group, organised in a 2x2 matrix with Factor A = ‘Presence of TPB components’, and Factor B = ‘Tailoring’. ........................................... 168
Table 5.14 NQuery output, based on feasibility data, showing the power obtained for main and interaction effects for three different possible sample sizes. ..................................................... 168
Table 6.1 Study groups. ........................................................................................................ 170
Table 6.2 The table shows when responses to the primary intention variable were considered as appropriate or inappropriate. ................................................................. 174
Table 6.3 Number and percentage of participants who were randomised and completed the study per group. ........................................................................................................ 177
Table 6.4 Demographic data of the sample (N = 212). .......................................................... 179
Table 6.5 Number and percentage of participants selecting each of the four TPB-statements (N=212). ............................................................................................................. 180
Table 6.6 Results for the Sheirer-Ray-Hare test using a randomly selected sample of n=50 per group (N=200). ........................................................................................................ 182
Table 6.7 Correlations between demographic variables and primary outcome intention....... 183
Table 6.8 Correlation between intention to seek help and presence of symptoms. ............. 183
Table 6.9 Regression model of demographic and symptom variables predicting intention to seek help ........................................................................................................................ 184
Table 6.10 Number of participants who completed the TPB questionnaire per study group. .. 185
Table 6.11 Results of the Sheirer-Ray-Hare test for the TPB questionnaire and lung cancer risk perception scale (N=120). ................................................................................... 185
Table 6.12 Frequency of selection of ‘Next steps’ options across the four study groups....... 187
Table 6.13 Sample sizes required to detect different effect sizes. ........................................ 195
Table 7.1 Study groups for Trial 2 ....................................................................................... 196
Table 7.2 User feedback obtained during the Think Aloud Evaluation. ........................................ 204
Table 7.3 Sample distribution across the four study groups. .................................................. 207
Table 7.4 Self-reported sample characteristics (N=253). ....................................................... 208
Table 7.5 Key words used to access the study link via the AdWords campaign. ..................... 209
Table 7.6 Mean self-reported likelihood of seeking help before and after treatment, and change from pre to post, across the four study groups. .......................................................... 210
Table 7.7 Mean self-reported likelihood of seeking help before and after treatment, for participants receiving tailored/untailored information, and TPB-component/no TPB-component. ............................................................................................................................................................. 212
Table 7.8 Summary results of the two-way mixed ANOVA on 20% trimmed means. ............ 212
Table 7.9 Observed and expected count of the number of participants responding appropriately/inappropriately to post-treatment self-reported likelihood of visiting a doctor, by factor level. ........................................................................................................................................ 212
Table 7.10 Change (post-pre) in self-reported likelihood of visiting a doctor for those aged ≥50 years ........................................................................................................................................... 213
Table 7.11 Summary results of the two-way mixed ANOVA on 20% trimmed means for those aged ≥50 years. .................................................................................................................. 214
Table 7.12 Change (post-pre) in self-reported likelihood of visiting a doctor, for those who were reportedly smokers or ex-smokers ...................................................................................... 215
Table 7.13 Summary results of the two-way mixed ANOVA on 20% trimmed means for self-reported smokers or ex-smokers only................................................................................................. 215
Table 7.14 Participants’ responses when asked, following treatment, what they would like to do next. Participants selected one of four options (n=252, 1 missing). .................................................. 216
Table 7.15 Results from the two-way ANOVA using 20% trimmed means on the secondary outcome variables. ...................................................................................................................... 217
List of Figures

Figure 1.1 Time intervals from first symptoms until treatment, from Weller et al. (2012). ........ 28
Figure 1.2 Pathways to treatment, from Walter et al. (2012). .................................................. 28
Figure 2.1 PRISMA diagram for the article identification process. .............................................. 41
Figure 3.1 Participant recruitment flow diagram (STROBE diagram). ........................................ 86
Figure 3.2 Proportion of the sample reporting various lengths of time between first noticing symptoms and contacting health services. ................................................................. 89
Figure 3.3 Use of lay information sources to appraise symptoms prior to diagnosis by participants and family/friends (“proxies”). ................................................................. 90
Figure 4.1 A screenshot of the NHS website on lung cancer symptoms (NHS choices, n.d.). 108
Figure 4.2 A screenshot of the cancer.org website on lung cancer symptoms (Cancer.org, n.d.). ......................................................................................................................... 109
Figure 4.3 MRC Guidance for developing and evaluating complex interventions from Craig et al., 2008 (2008). ................................................................. 109
Figure 4.4 Conceptual framework of the Theory of Planned Behaviour, from (Scottish Government, 2008). ................................................................. 115
Figure 4.5 Screenshot showing how I targeted behavioural beliefs on the intervention website (version 1). ................................................................. 124
Figure 4.6 Screenshot showing how I targeted normative beliefs about family/friends on the intervention website (version 1). ................................................................. 125
Figure 4.7 Screenshot showing how I targeted normative beliefs about health professionals on the intervention website (version 1). ................................................................. 125
Figure 4.8 Screenshot showing how I targeted control beliefs on the intervention website (version 1). ................................................................. 126
Figure 4.9 The screenshot shows an example of tailored symptom information a user would receive who endorsed a cough and unintended weight loss during the initial survey (website version 1). ................................................................. 127
Figure 4.10 The screenshot shows an example of tailored risk factor information a user would receive who reported being an ex-smoker, and age above 40 (website version 1). .............. 127
Figure 4.11 Screenshot of the tailored intervention summary (website version 1). ............... 128
Figure 4.12 Structure of the version of the intervention implemented in the feasibility study... 136
Figure 5.1 This flowchart details the study procedure. TPB = Theory of Planned Behaviour. INT = Intervention group, CG-TPB = Comparison group 1: untailored + TPB components, CG-Tail = Comparison group 2: tailored, no TPB components, UC = usual care, based on Roy Castle website on lung cancer symptoms [6].

Figure 5.2 Illustration of cases in which button click behaviour is considered appropriate or inappropriate.

Figure 5.3 Age distribution of visitors of the study website based on estimates by Google Analytics.

Figure 5.4 Number of participants per study group, and proportion (95% CI) out of the total sample size (N=130).

Figure 5.5 Number of participants reporting various symptoms.

Figure 5.6 Line graph showing the mean (95% CI) intention to seek medical help (7-point scale) of the four study groups.

Figure 5.7 The information on NICE clinical guidelines was visually highlighted for the full trial, to ensure this information would be read by users, as this information could be useful to individuals who have already sought help.

Figure 6.1 Study procedure. TPB = Theory of Planned Behaviour. INT = Intervention group, CG-TPB = Comparison group 1: untailored + TPB-components, CG-Tail = Comparison group 2: tailored, no TPB-components, CG-none= Comparison group 3: untailored, no TPB-components. Changes from the feasibility study to the trial are highlighted in green font.

Figure 6.2 Number of participants reporting each symptom, and proportion reporting a duration of ≥3 weeks, or less.

Figure 6.3 Mean (95% CI) intention to see a GP about the symptoms (measured on 7-point scale) across the four study groups.

Figure 6.4 Mean (95% CI) intention to see a GP about the symptoms among participants who received TPB-components, compared to those who received no TPB-components.

Figure 6.5 Mean (95% CI) intention to see a GP about the symptoms among participants who received tailored information, compared to those who received untailored information.

Figure 6.6 Distribution of standardised residuals for the primary outcome variable, intention to see a GP about the symptoms.

Figure 6.7 Mean (95% CIs) intention by affected individuals and proxies.
Figure 7.1 TPB-component used for Trial 2. Sub-components were shown to participants consecutively; participants clicked through these using the arrows and were only able to proceed with the study once all sub-components had been viewed. ........................................ 198

Figure 7.2 This flowchart details the study procedure. TPB = Theory of Planned Behaviour. INT = Intervention group, CG-TPB = Comparison group 1: untailored + TPB-components, CG-Tail = Comparison group 2: tailored, no TPB-components, CG-none = Comparison group 3: untailored, no TPB-components. Changes from the previous trial are shown in green font. ... 201

Figure 7.3 Symptom distribution across the sample (N=253). ......................................................... 208

Figure 7.4 Mean (95% CI) change in self-reported likelihood of visiting a doctor from pre to post across the four study groups............................................................... 211

Figure 7.5 Mean (95% CI) reported likelihood of seeking help at pre and post treatment for those who received tailored and those who received untailored information, for those aged ≥50. ................................................................................................................................. 214

Figure 7.6 Mean (95% CI) reported likelihood of seeking help at pre and post treatment for those who received the TPB-component and those who did not, for those aged ≥50. .......... 214
Glossary and acronyms

CI: Confidence interval

**Feasibility study**: Feasibility work intends to assess whether planned study procedures are viable, practicable and workable, taking into account clinical, cultural, logistic, economic and ethical factors (Porta, 2008). Feasibility studies also often provide estimates of parameters needed to design the main study, e.g. for sample size calculations.

GP: General Practitioner

**Help-seeking**: Help-seeking is the process of perceiving, interpreting and appraising symptoms, deciding whether to seek professional help, and presenting (or not presenting) to health care services (Scott & Walter, 2010). This process is not limited to the time from symptom onset to first presentation to health services, because several cycles of appraisal and re-appraisal of symptoms and subsequent decisions of whether to (re-) present to health services can take place until a final diagnosis is made. For the purpose of this thesis help-seeking behaviour spans the time from first symptom perception until diagnosis.

Internet: The Internet is a system of interconnected networks which connects computers across the globe using the standard Internet protocol suite. The Internet is an infrastructure used to access other services deployed on top of it, including e-mail, instant messaging and the World Wide Web, commonly referred to as the Web.

**Pilot study**: A pilot study is a study which employs, on a small scale, the same methods as a planned, larger-scale trial, in order to test whether the planned components can all work together smoothly (Porta, 2008).

PPI: Patient and Public Involvement

**Proxies**: The term “proxy” is used throughout the thesis to denote family/friends of patients.

RCT: Randomised controlled trial

**Think Aloud**: In the Think Aloud paradigm (also known as ‘cognitive interviewing’), users voice their thoughts aloud and vocalise their actions while performing a task, thus helping the researcher to understand the users’ actions and identify issues as well as useful features (Krug, 2013; Lewis, 1982).

TPB: Theory of Planned Behaviour.

Web: The Web is a collection of HyperText Markup Language (HTML) documents, or ‘Web pages’. These can be accessed using a Web browser such as Internet Explorer, Google Chrome or Mozilla Firefox (Webopedia, 2010; Whitehead & Maran, 1997).
Abstract

Background. Lung cancer is the leading cause of cancer deaths worldwide. Low survival rates have been attributed to delays to diagnosis, and some patients report having symptoms for several months before presenting to health services. Strategies are needed to encourage timely help-seeking. The Web is increasingly used as a health information source.

Aim. The aim of this thesis is to explore whether the Web plays a role in help-seeking behaviour of people with lung cancer prior to diagnosis, and how the Web can be utilised to encourage earlier presentation to health services for symptomatic people.

Systematic review. To begin, I carried out a systematic review of the literature (N=34), which highlighted a scarcity of research on Web use for symptom appraisal among cancer populations.

Mixed-methods study. I conducted a survey with recently diagnosed (≤ 6 months) lung cancer patients (N=113). Based on survey responses, I purposively selected a sub-sample of patients and their family/friends (“proxies”) for semi-structured interviews (N=33). In the survey, 20.4% of participants reported they or proxies had researched their condition online before the diagnosis. Interview results suggest perceived impacts of online information on symptom appraisal, forming the decision to seek help, and on interactions with healthcare professionals.

Intervention development and evaluation. Based on my findings, I developed a Web-based intervention. The intervention provides tailored information about lung cancer, and uses components based on the Theory of Planned Behaviour (“TPB-components”) to encourage earlier help-seeking. This intervention was tested in an online feasibility study (N=130), and subsequently in an online randomised controlled trial (N=212) with a factorial design to test main and interaction effects of tailoring and TPB-components. The feasibility study and first trial helped identify methodological issues which were addressed in a second trial with a mixed factorial design. This trial (N=253) indicated that the self-reported likelihood of visiting a doctor increased significantly by 11.8% from before to after viewing study information (p<0.001), but no effects of tailoring or TPB-components were found. When examining only those aged ≥50 years, who are at highest risk of lung cancer, those receiving tailored information reported an increase of 13.2% in likelihood of seeking help, compared to 3.2% in the untailored group (p=0.01). Participants aged 50+ receiving the TPB-component reported a larger increase (13.8%) than those who did not receive the TPB-component (5.2%), but this did not meet the significance criterion (p=0.054).

Conclusions. According to patients’ perceptions, the Web can impact on processes in the appraisal, help-seeking and diagnostic interval leading up to diagnosis. Presentation of information about symptoms and risk factors online can produce significant effects on self-reported likelihood of seeking help. These findings indicate that there is potential for the Web to be utilised in the endeavour to educate the public about symptoms and to promote earlier presentation to health services, but due to small effects and differential dropout in this study, further research is required.
Lay Abstract

The aim of this PhD thesis was to explore whether people with lung cancer use the Internet to look up their symptoms before they are diagnosed, and whether this influences their decision to see a doctor. This is important because previous research has shown that people with lung cancer sometimes wait for a long time before seeing a doctor. By the time the disease is diagnosed, it is often very difficult to treat.

To start, I reviewed previous research on using the Web to appraise symptoms, using a systematic approach to ensure I found all relevant articles. This showed me that there is very little research on this topic. I then conducted a survey with recently diagnosed lung cancer patients, and interviewed some patients and their family members in more depth. About a fifth of the patients I surveyed said that either they, or one of their family members or friends, had looked up their symptoms online before they were diagnosed. In the interviews, participants said that the information they found online influenced how they appraised symptoms, their decision to see a doctor and how they communicated with their doctors.

This showed me that the Internet can play a role in help-seeking for lung cancer symptoms. Next, I investigated different ways of presenting information, to see if we can improve online information to make it more helpful. I developed a website which provides people with personalised information about lung cancer, depending on their unique symptoms and risk factors. The website also includes some components designed to change beliefs that might stop people from seeking help (I call these “TPB-components” because they are based on a psychological theory called the Theory of Planned Behaviour). The TPB-components consist of quotes from health professionals and fictional patients or family members, and a list of steps to show people how they can seek help.

To investigate if the website encourages people to seek help for their symptoms, I conducted an online study with people who reported having symptoms that might be related to lung cancer. Participants were randomly assigned to different study groups which I then compared. To test the effects of information personalisation and the TPB-components, participants received different combinations of these two features – either both, one of them, or neither. Participants were asked to state their intention to seek help for the symptoms before and after viewing the information.

My results showed that intention to seek help increased in all participants, regardless of how the information was presented to them. When I looked specifically at those aged 50+ years, the results suggest that people who saw personalised information had bigger increases in their intention to seek help than people who saw generic information.

Overall, the findings of this thesis show that the Internet does play a role in people’s help-seeking for lung cancer symptoms. They also suggest that we might be able to present online information differently to encourage more people to see their doctor promptly, but further research is needed.
Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DoculInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University’s policy on Presentation of Theses.
Chapters 2 and 3 contain material published in form of open-access articles distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The material presented in Chapter 2 is published in (Mueller, Jay, Harper, Davies, et al., 2017) and is co-authored by my supervisors and two other researchers. I designed the review protocol, conducted the searches, assessed the studies for eligibility, extracted, synthesised and interpreted the data, and subsequently wrote the review. My supervisors provided advice, guidance and critical appraisal of my work. The two remaining authors (JV and AD) were included to assess studies for eligibility independently of myself in order to enhance rigour and ensure the criteria for a systematic review are met. The material presented in chapter 3 is published in (Mueller, Jay, Harper, & Todd, 2017) and is co-authored by my supervisors, who provided advice, guidance and critical appraisal of my work. Under this supervision, I designed the study, collected, analysed and interpreted the data, and drafted the written work.
Acknowledgements

First and foremost, I want to thank my supervisors Professor Chris Todd, Dr Caroline Jay and Dr Simon Harper for their support and guidance over the past four years. I am extremely grateful and it has been a privilege to learn and grow under your supervision. A very special thank you to my main supervisor – thank you Chris for always sensing when I was on the verge of a freak out, and helping to put things into perspective.

I would also like to thank Dr Morag Farquhar, who helped with my supervision during my first year while Chris was away. I really appreciate the constructive feedback and guidance you provided.

I also want to thank the consultants, nurses and general practitioners who facilitated access to the clinics, and who provided clinical input for the development of the intervention. A special thank you to Dr Fiona Blackhall, Dr Amelie Harle, Dr Yvonne Summers and Dr Liam Hosie. Thank you also to the participants who provided data for the studies.

My sincere gratitude extends to the statisticians who provided advice for my data analysis and sample size calculations.

I thank the Medical Research Council for funding my PhD with a Doctoral Training Partnerships grant, and the University of Manchester, for supplementing my MRC stipend with a President’s Doctoral Scholar Award.

To the members of the Interaction Analysis and Modelling laboratory – thank you all so much, it has been a real pleasure to work with and alongside you all (even you Kevin). A very special thank you to Alan Davies, for programming the When2Go website. Thank you for all the hard work and long hours that you put in. I could not have done any of this without your help. Also, thank you to Julio Vega, for helping us host the website on the server. You are an absolute star. And always remember that in Linux, image file extensions are in CAPITALS.

Thank you to my friends, Katie, Lora, Emily, Grace – thank you girls for helping to keep me sane throughout these four years. You were always there with advice (and wine!) when I got stuck on some PhD work.

Nick, thank you for always being there for me. I could have done it without you – but I’m really glad I didn’t have to!

I also want to thank my family. I would not be who or where I am today without your continued support and encouragement.

Finally, I want to thank Dr Katrin Seeher and Dr Arpana Verma, who have had a great impact on my academic pursuits.

If I have forgotten anyone, I’m really sorry. As C. Jenc once said, “I am burning on a low flame,” and my head is full of thesis.
The Author

I completed a BSc in Psychology at the Georg-August University in Goettingen, Germany in 2011. My bachelor’s thesis involved the evaluation of an Internet-based self-help training for recurrent headache in childhood and adolescence, through which I became interested in the use of Internet-based technology in health promotion. During my undergraduate degree I also completed a three-month research internship at the Dementia Collaborative Research Centre at the University of New South Wales, Australia, which determined my decision to pursue a career in research. This was followed by an MSc in Health Psychology at St Andrews University. My MSc thesis focused on sexual health behaviour and used structural equation modelling to assess predictors of condom use among adolescents. I began my doctoral research in 2013, and have also worked part-time as a research assistant in public health for the past two years.
Chapter 1: Introduction

1.1 Overview

In this thesis I explore how people with lung cancer use the Web as a health information source to appraise their condition prior to diagnosis, and whether the Web plays a role in their help-seeking behaviour. I also explore how the Web can be utilised to encourage earlier presentation to health services and thus promote earlier diagnosis, thereby facilitating treatment.

In this first chapter I present background information on key topics, demonstrating why it is important to investigate the role of the Web in the time before a lung cancer diagnosis. In Chapter two I then go on to present findings from a systematic review on Web use for symptom appraisal. Subsequently I describe an exploratory study I conducted with recently diagnosed lung cancer patients and their family/friends (“proxies”), to explore their subjective experiences of the time before diagnosis and use of the Web. I then portray how I used findings from the literature review and the first study to develop a Web-based intervention to encourage early help-seeking for lung cancer symptoms, and how I developed and undertook a study to assess its effectiveness in an online randomised controlled trial. Finally I draw these threads together in discussion.

Lung cancer is the leading cause of cancer deaths worldwide (Fitzmaurice et al., 2015). It accounts for approximately 19% of all cancers and 40,000 deaths per year in the UK (Longmore, Wilkinson, Baldwin, & Wallin, 2014). Lung cancer survival rates in the UK are lower than in other European countries which have comparable per capita total national expenditure for health (De Angelis et al., 2014; Holmberg et al., 2010), with 5-year age-standardised survival rates of 9% (95% CI: 8.8-9.1) in the UK while the European average lies at 13% (95% CI 12.9-13.1) (De Angelis et al., 2014). Research suggests these low survival rates are attributable to delays to diagnosis (Holmberg et al., 2010). Times to diagnosis and treatment in lung cancer often exceed recommendations (Olsson, Schultz, & Gould, 2009), and studies have documented that some patients have symptoms for several months before presenting to health services (Corner, Hopkinson, Fitzsimmons, Barclay, & Muers, 2005; Corner, Hopkinson, & Roffe, 2006; Tod, Craven, & Allmark, 2008). Delays to diagnosis can lead to advanced stage of disease when treatment is commenced (Allgar & Neal, 2005; Radzikowska, Roszkowski-Sliz, & Glaz, 2012), which decreases treatability and chance of survival (Zee, Eisen, & Carney, 2008).
It is estimated that every year up to 7000 cancer deaths occurring within five years of diagnosis could be avoided if the UK matched European mean survival rates (Richards, 2009b), which includes 900 lung cancer deaths (Abdel-Rahman, Stockton, Rachet, Hakulinen, & Coleman, 2009). This introductory chapter will examine pathways to diagnosis to lung cancer, with particular emphasis on the role of patients, symptom appraisal and help-seeking behaviour.

Lung cancer is broadly categorised into small-cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC), which refers to histological differences of the tumour. In SCLC, metastasis proceeds more rapidly and widely than in NSCLC, leading to lower survival rates. NSCLC incorporates the three sub-categories adenocarcinoma, squamous cell carcinoma and large cell undifferentiated carcinoma. In both NSCLC and SCLC tumours grow very rapidly, and are commonly detected at a late stage with metastatic spread (Dilworth & Baldwin, 2001). Without metastatic spread, tumours can be treated through surgery. At advanced disease stages, surgery is not possible and combinations of chemotherapy and radiotherapy are commonly employed (Longmore et al., 2014). Treatment effectiveness and prognosis depend largely on the stage of the tumour at diagnosis.

Over the past two decades, our understanding of lung cancer risk factors has developed considerably. The most important predictor of lung cancer is smoking status; approx. 90% of lung cancer cases are caused by smoking, and the risk of lung cancer is directly related to the number of cigarettes smoked per day (Bourke & Burns, 2011). Other risk factors include older age, genetic factors, exposure to ionising radiation such as radon gas which can arise from building material, occupational exposure to asbestos, and having a history of other cancers involving the lung, head or neck (Bourke & Burns, 2011; Dilworth & Baldwin, 2001).

Having described the relationship between delayed diagnosis in lung cancer and survival rates, I will now explore in more detail which processes underlie the time from disease onset to diagnosis, and explain why this thesis focuses on patient factors.

1.2 Why focus on patient factors in lung cancer?

A substantial amount of research around early diagnosis of lung cancer has been devoted to healthcare professionals and system factors, and several interventions are underway to reduce unnecessary delays occurring at these points. For example, a range of support is available to
help General Practitioners (GPs) assess patients effectively and identify cancer warning signs at an early stage, such as risk assessment and decision support tools (Department of Health, 2011b). In terms of healthcare system factors, waiting time standards appear to have reduced system-related delays (Department of Health, 2011a).

Patient factors, on the other hand, are more difficult to target because the group of people who do not present to health services is not clearly identified and cannot be addressed directly. In addition to this, there is to date no national screening programme for the early detection of lung cancer in the UK because evidence for the effectiveness of screening has been mixed (Duffy, Field, Allgood, & Seigneurin, 2014; Gill et al., 2013; Prosch, Mostbeck, Latshang, Lo Cascio, & Russi, 2013). Until a robust evidence base exists and a programme can be implemented in the UK, diagnosis relies primarily on patients detecting their own symptoms and presenting to health services promptly. Even if a screening programme is implemented in the UK in future, early diagnosis will still rely heavily on recognition and presentation of warning signs by patients, and efforts are needed to promote symptom awareness (Richards, 2009a). Moreover, evidence suggests the time taken by patients to seek help has a stronger impact on survival than practitioner or healthcare factors (C. S. Thompson & Hetzel, 1997) and that efforts to improve earlier presentation of patients could have considerable impacts on survival outcomes (Alwan, Howells, Skinner, Munavvar, & Singh, 2010; Hassan, 2010).

Research has indicated that patients in the UK may experience symptoms for several months before presenting to health services (Corner et al., 2005, 2006; Tod et al., 2008). In addition to this, awareness of lung cancer symptoms among the public in the UK appears to be low (Simon et al., 2012), and perceived barriers to early presentation, such as worry about wasting the doctor’s time, are higher in the UK than in other high-income countries (Forbes et al., 2013). On average, people can only recall one symptom, and many symptoms such as shoulder pain or changes in an existing cough are largely unknown to the population (Simon et al., 2012). The public also seems to be unaware of many risk factors for lung cancer, such as having a history of other lung diseases or exposure to air pollution, and even smokers are often unrealistically optimistic about their personal elevated risk (Arnett, 2000). Self-detection of lung cancer symptoms is particularly difficult because symptoms tend to be vague and ambiguous: Common symptoms include a cough, breathlessness, fatigue, and loss of appetite or weight loss (Bourke & Burns, 2011; Dilworth & Baldwin, 2001). Clear warning signs such as haemoptysis are
typically only experienced by a small proportion of patients (Walter et al., 2015). Symptoms are therefore often minimised by those afflicted and attributed to less serious causes such as a common cold or not recognised as symptoms as all (Corner et al., 2006). Those at highest risk for lung cancer tend to be older and (ex-)smokers, and therefore often have existing comorbidities such as chronic obstructive pulmonary disease (COPD) which can mask the symptoms (Tod & Joanne, 2010; Walton et al., 2013) and lead to a longer time to diagnosis (Walter et al., 2015). Furthermore, symptoms tend to develop gradually and slowly over time, so that those afflicted often grow accustomed to these physical changes and normalise them as a result (Brindle, Pope, Corner, Leydon, & Banerjee, 2013). Moreover patients are reluctant to endorse symptoms when they are phrased in a manner that implies ill health or disease but more readily admit to experiencing the symptoms when they are phrased in a less threatening manner, such as “discomfort” instead of “pain” (Brindle et al., 2013). This demonstrates that symptoms are often not perceived as indicative of a specific disease but are associated with normal or minor health changes that are expected to resolve on their own.

Fear and fatalistic beliefs may also play a role in individuals’ reluctance to seek medical advice (Corner et al., 2005, 2006; McCutchan, Wood, Edwards, Richards, & Brain, 2015; McLachlan et al., 2015; Shim, Brindle, Simon, & George, 2014; L. K. Smith, Pope, & Botha, 2005; Tod & Joanne, 2010). Often individuals hold fatalistic beliefs about cancer and its treatability (Walton et al., 2013). Furthermore people may be worried about the stigma and blame attached to lung cancer as this disease is linked to unhealthy lifestyles (L. K. Smith et al., 2005; Tod & Joanne, 2010). Delay to presenting to health services can also be increased if people are worried about bothering doctors with their problems (Robb et al., 2009; L. K. Smith et al., 2005; Tod & Joanne, 2010).

Thus, although practitioner and health system-related factors are clearly important determinants in the delay to diagnosis observed in lung cancer patients, patient factors in the time from symptom onset until diagnosis play a considerable role and need to be tackled. For this reason, this research project focuses on patient-related factors in the pathway to diagnosis of lung cancer.
1.3 Help-seeking behaviour

Above, I argue why this project focuses on patient-related factors and intervals in the diagnostic process. I will now examine patients’ roles in further detail by discussing previous literature around patients’ help-seeking behaviour, and interventions which have aimed to target this.

1.3.1 What is ‘help-seeking behaviour’?

There is a rich tradition in psychological and social sciences of conceptualising help-seeking behaviour and exploring which factors contribute to health care utilisation. Help-seeking is the process of perceiving, interpreting and appraising symptoms, deciding whether to seek professional help, and presenting (or not presenting) to health care services (Scott & Walter, 2010). It is a “complex decision-making process instigated by a problem that challenges personal abilities” (Cornally & McCarthy, 2011). It is characterised as a problem-focused, intentional, planned behaviour and interpersonal interaction with a healthcare professional (Cornally & McCarthy, 2011). Because symptoms are not always clear and distinct and not all physical changes relate to illness, this process tends to be complex and depends on a number of factors and circumstances.

A comprehensive, parsimonious overview of processes involved in symptom appraisal and help-seeking behaviour is provided by the Psychophysiological Comparison Theory (B. L. Andersen & Cacioppo, 1995), which combines insights from various theories and publications focused on illness appraisal and help-seeking among cancer patients. The theory assumes that individuals are motivated to understand bodily changes, but there are several barriers to attributing changes to illness. It postulates that the more unexpected, salient, personally relevant the symptoms and/or the higher the perceived consequences, the higher the motivation to understand the symptoms. The theory also suggests that individuals undertake psychophysiological comparisons by evaluating symptoms in the context of external circumstances (e.g. smoky or dusty work environments) or symptoms of illnesses the person is familiar with. Should individuals consider more serious causes for their symptoms, the theory postulates that they will be optimistically biased towards assuming threatening illnesses are more likely to happen to others than to themselves.

Mood and cognitions also play a crucial role in symptom perception. For example, stress and anxiety may exacerbate symptoms while relaxation may improve them (Ogden, 2012). Not only
the symptom itself but also the impact this symptom has on everyday life will influence the
decision to seek help: symptoms often trigger help-seeking if the individual perceives them as
disruptive to their work, physical activity or social relations (Hay, 2008; Zola, 1973). Help-
seeking is also socially triggered when other people encourage the individual to see a doctor (B.
L. Andersen & Cacioppo, 1995; Rosenstock, Strecher, & Becker, 1988).

Individuals also tend to weigh up the costs and benefits of seeking professional care before
forming a decision, considering whether effective treatment is available, whether this will involve
negative side effects, and whether having their problem examined will be embarrassing or
uncomfortable (Ogden, 2012). Furthermore, help-seeking depends on whether an individual
wishes to take on the benefits and obligations of the ‘sick role’, such as being absolved from
chores and duties but having to comply with physician advice (Parsons, 1951). Cultural
assumptions can impact on illness perception and behaviour by influencing values, meanings
and beliefs associated with illnesses and by determining common practices. Mechanic (1978)
argues that cultural assumptions are key to forming health decisions, and can influence, for
example, whether a symptom is perceived as ‘normal’ or requiring medical attention.

Previously, the time between first symptom perception and presentation to health services was
termed ‘patient delay’, though newer recommendations suggest using the more neutral term
‘patient interval’ (Weller et al., 2012). Pack and Gallo (1938), who first discussed the term
‘patient delay’ in relation to cancer in 1938, described initial patient delay as the time that
elapses between first symptom perception to first visiting a physician, and subsequent patient
delay as delay occurring after a physician is consulted. The authors arbitrarily defined a delay
period under three months as “reasonable” while considering over 3 months to be “undue”
(Pack & Gallo, 1938, p. 443).

Safer et al. (1979) subsequently developed a model which describes patient delay in three
stages: 1) Appraisal delay (occurs when a person notices physical changes but does not realise
immediately that these are symptoms of an illness), 2) Illness delay (occurs when individuals
realise that they are ill but delay consulting a health professional, and 3) Utilisation delay
(occurs when individuals have made the decision to seek medical help but delay in utilising the
service).
Building on Safer et al.’s (1979) work, Andersen and Cacioppo’s (1995) model of patient delay adopts the stages of appraisal and illness delay but describes the process of presenting to health services in three further stages. From the moment the decision to seek medical help is formed, a person can delay in making an appointment, e.g. by putting off making the phone call (termed *behavioural delay*). Subsequently, *scheduling delay* can occur e.g. if the person makes the phone call, but asks for a later appointment even if an earlier one is available to accommodate other obligations or due to emotional factors such as fear. Finally, *treatment delay* can occur between the first medical consultation and commencement of treatment. The final three stages can also be influenced by practitioner and system delay, e.g. due to limited appointments.

In line with this concept of categorising the time leading up to diagnosis into intervals, Weller et al. (2012) adapted a model proposed by Olesen et al. (2009) to illustrate milestones and time intervals from first symptoms until commencement of treatment (Figure 1.1). In this model, the ‘patient interval’ describes the time from first noticing bodily changes to first presentation or clinical appearance. Thereafter, the diagnostic interval commences, which includes both primary care and secondary care. The primary care interval encompasses first presentation, first investigation and first referral to secondary care. The ensuing secondary care interval spans the time from first specialist visit until diagnosis, and subsequently the treatment interval commences. This model provides a more parsimonious account of time intervals in the diagnostic pathway than previous approaches, which described only patient, health professional and system delay (Burgess, Ramirez, Richards, & Love, 1998). A further improvement of this model over previous approaches is the replacement of the term ‘delay’, which suggests a certain assignment of blame, with the more neutral term ‘interval’.

What the above model fails to convey, however, are loops and overlaps between different intervals. The model suggests a linear progression from first symptoms to treatment, which does not always occur. Patients may, for example, present to primary care multiple times before an investigation of symptoms takes place. Furthermore, patient factors may be relevant even during primary care and secondary care intervals, for example if patients fail to attend appointments.
A model of pathways to treatment developed as an extension of the earlier model proposed by Andersen et al. (1995) provides a less linear, more iterative representation (Scott, Walter, Webster, Sutton, & Emery, 2013; Walter, Webster, Scott, & Emery, 2012) (Figure 1.2). This model acknowledges that patient factors, healthcare provider, and system factors as well as disease factors can influence all intervals of the pathway to treatment. It also illustrates the loops between different stages. The *appraisal interval* includes an interpretation of bodily changes (given the event that these changes are detected). The *help-seeking interval* describes the time taken to act upon these interpretations (Weller et al., 2012). The *diagnostic interval* involves appraisal by a healthcare provider, investigations, referrals and appointments, and, in the event of a diagnosis, leads to the start of treatment.

To summarise, help-seeking behaviour depends on the nature of symptoms, symptom perceptions (physical and emotional, and their impact on everyday life), cost-benefit assessments of help-seeking behaviour, influence of significant others, and cultural assumptions. The above models demonstrate the complex nature of the diagnostic process, with interplays between factors related to the patient, healthcare providers and the healthcare system. For the reasons outlined above, I do not adopt the division of the period from symptom onset to diagnosis into separate intervals affected by only one group of actors somewhat artificial. Therefore, I examine the entire period from symptom perception to diagnosis, with particular focus on patient factors.

Thus, for the purpose of this study, help-seeking is defined as the process of perceiving, interpreting and appraising symptoms, deciding whether to seek professional help, and presenting (or not presenting) to health care services. Due to the cyclical, non-linear progression of events leading up to diagnosis, “help-seeking” in this thesis encompasses the time period from first symptom perception until final diagnosis.
1.3.2 Measuring help-seeking

Despite the considerable amount of research and number of theoretical frameworks devoted to help-seeking behaviour, measuring help-seeking accurately remains a challenge (R. Andersen, Vedsted, Olesen, Bro, & Sondergaard, 2009; Scott & Walter, 2010). To date, there is no validated, reliable and accurate measure of help-seeking. One of the main limitations is that help-seeking intervals are typically measured retrospectively, and patients often do not
accurately remember the events leading up to diagnosis (Hess et al., 2012; Sheppard, Kumar, Buckley, Shaw, & Raza, 2008). Furthermore, no clear consensus exists on terms and definitions (Scott & Walter, 2010), resulting in a wide range of different methods used in this field (Allgar & Neal, 2005), which hampers comparability of findings. For example, previous approaches have involved extracting information from patient records (which are often incomplete or poorly recorded), use of questionnaires and interviews, or a triangulation of these (Scott & Walter, 2010).

1.3.3 Help-seeking and cancer

Help-seeking has been extensively researched in the specific context of cancer. Many studies have used qualitative methods to explore patients’ experiences prior to diagnosis (de Nooijer, Lechner, & de Vries, 2001; De Nooijer, Lechner, & De Vries, 2001; Molassiotis, Wilson, Brunton, & Chandler, 2010; L. K. Smith et al., 2005), and some have used quantitative survey methods to assess help-seeking across larger samples (Hansen, Olesen, Sorensen, Sokolowski, & Sondergaard, 2008; Mor, Masterson-Allen, Goldberg, Guadagnoli, & Wool, 1990; Niksic et al., 2015). Across different cancer types, facilitators of help-seeking include symptom recognition and attribution of symptoms to cancer, with specific, widely-known symptoms like lumps more likely to elicit help-seeking than vague, ambiguous symptoms like coughing (de Nooijer et al., 2001; De Nooijer et al., 2001; L. K. Smith et al., 2005). Support from other people has also emerged as an important facilitator across different cancer types (L. K. Smith et al., 2005). Barriers include fear of cancer (in terms of pain, suffering or death), worry about wasting the doctor’s time, fear of embarrassment for sensitive or stigmatised symptoms, and misattribution of symptoms to existing or less serious conditions (de Nooijer et al., 2001; De Nooijer et al., 2001; Mor et al., 1990; L. K. Smith et al., 2005). These findings are mirrored in studies on help-seeking among lung cancer patients, showing that these facilitators and barriers play important roles across different cancer types (Birt et al., 2014; Brindle et al., 2013; Caswell et al., 2017; Corner et al., 2005, 2006; Levealahti, Tishelman, & Ohlen, 2007; Tod et al., 2008; Tod & Joanne, 2010; Walton et al., 2013).

Thus far, I have described the pathway to diagnosis in lung cancer, the role that patients play in this process, and the factors that may impact on it. The next logical question is: what can be done to reduce the length of the time interval from first symptom perception until diagnosis? In
the following section, I will discuss efforts which have been undertaken to date to target patient factors in the pathway to diagnosis in (lung) cancer, and I will show that a viable and to date under-utilised option lies in the Web as a health information source.

1.3.4 Previous interventions for help-seeking behaviour in cancer

The National Awareness and Early Diagnosis Initiative was launched in England in 2008 as part of the government’s strategy to improve cancer outcomes (Richards, 2009a). The aim of this initiative is to support and drive forward early diagnosis of cancer. This includes improving early diagnosis in patients who present with symptoms (as opposed to those identified via screening). A key step towards this aim is raising awareness of signs and symptoms among the public, and targeting negative beliefs about cancer (Richards, 2009a). Mass media campaigns have successfully raised the awareness of warning signs in lung cancer (Ironmonger et al., 2014), cervical and breast cancer (Jenkins et al., 1999), oral cancer (Eadie, MacKintosh, MacAskill, & Brown, 2009), and skin and colon cancer (Broadwater C, Heins J, Hoelscher C, & Mangone A, 2004). In the UK, a national NHS (National Health Service) campaign to increase cancer awareness under the brand name “Be Clear on Cancer” has had several positive effects on screening uptake and symptom awareness (Peacock, Clayton, Atkinson, Tierney, & Lund, 2013). Interventions are typically delivered in a community-based format (i.e. information is disseminated to the community via events at pubs, clubs, mosques, supermarkets etc.), via mass-media (television or radio) or by healthcare providers in clinical settings (Richards, 2009a).

In the UK, lung cancer awareness raising campaigns have included ‘I’ll tackle it soon’, ‘3 Week Cough’, ‘Be Clear on Cancer’, and ‘Detect Cancer Early’ (Gordon, Magee, Jones, Phillipson, & Barrie, 2012). These campaigns were aimed at the broad public and attempted to disseminate clear, simple messages mostly addressing the cough symptom. Although these campaigns show some positive effects, evidence on their effectiveness tends to be limited and low in quality (Gordon et al., 2012). In a recent randomised controlled trial, Smith et al. (2013) evaluated the effectiveness of a theory-based behavioural primary-care intervention to reduce patient delay in lung cancer targeted towards those most at risk (smokers and former smokers ≥ 55 years). This intervention comprised a nurse consultation and showed some promising results in terms of intention to seek help and increases in consultation rates. However, implementation
of this intervention involves considerable investment of effort and resources, as high-risk individuals need to be identified at primary care centres and guided through the self-help manual by a trained nurse.

Overall, previous campaigns have focused on short, persuasive messages and have targeted broad audiences, or have involved healthcare provider-delivered material (i.e. opportunistic health promotion). Information is provided to individuals as passive recipients of information, i.e. individuals are likely to encounter the information when they are not actively seeking it, e.g. via a radio ad. While this approach is important, less has been done to target information sources people use to actively seek out information on cancer prior to diagnosis, such as the Web. This could have important implications, as sources accessed prior to diagnosis will likely have a direct impact on the decision to present, especially if sources are accessed in order to appraise existing symptoms.

1.4 The Internet, the World Wide Web and Health

Considering the rapidly increasing volume of health information on the Web and the increasing tendency of individuals to seek health information online, the Web could potentially be a key factor influencing awareness and health decisions (Andreassen et al., 2007; Chou, Liu, Post, & Hesse, 2011; Fox & Duggan, 2013). There is a growing body of literature which documents that individuals often research their symptoms online prior to presentation to health services, and that this impacts on decisions to seek help (Mueller, Jay, Harper, Davies, et al., 2017). Thus, targeting online health information could influence cancer symptom awareness and presentation to health services.

Before expanding further on health-related Web use, some technical terms need to be clarified. The terms ‘Internet’ and ‘Web’ are often used interchangeably although they are distinct. The Internet is a system of interconnected networks which connects computers across the globe using the standard Internet protocol suite. The Internet is an infrastructure used to access other services deployed on top of it, including e-mail, instant messaging and the World Wide Web, commonly referred to as the Web. The Web is a collection of HyperText Markup Language (HTML) documents, or ‘Web pages’. These can be accessed using a Web browser such as Internet Explorer, Google Chrome or Mozilla Firefox (Webopedia, 2010; Whitehead & Maran, 1997). The origins of the Internet date back to the early 1960s, when US and UK researchers...
first developed ideas to globally interconnect computers. The Web was created thirty years later, in the early 1990s, by the European Laboratory for Particle Physics as a means of sharing project information between researchers and facilitating cooperation. In 1993, the Internet became accessible to home users and the first publicly accessible website was created (Whitehead & Maran, 1997). Since then, Internet use has increased exponentially across the globe. According to world Internet usage statistics from 2017, 51.7% of the world’s population currently has access to the Internet and over 80.2% of the European population access the Internet (Internet World Stats, 2017).

Over recent years, the Web has become increasingly popular as a source for health-related information. Web use for health-related purposes has increased significantly, while the use of traditional health information sources has decreased or remained the same (Kummervold et al., 2008).

The Web as a source for health information has several benefits that highlight its potential to inform and influence peoples’ decisions about health care. It offers round-the-clock, anonymous and widespread access to information and also allows features such as interactivity (e.g. in forum discussions) and information tailoring to individual needs (Cline & Haynes, 2001; Ziebland et al., 2004). It can deliver specific content while also offering a large amount and range of information that is often controlled by professional bodies such as the NHS (Kirschning & von Kardorff, 2008). Research suggests online health information may even offer users greater autonomy and relatedness (feeling a psychological connection with relevant others) than face-to-face consultations with healthcare providers (S. T. Lee & Lin, 2016). While this raises concerns about the displacement of professional advice through online health information, which is often not scientific or peer-reviewed, evidence from the literature consistently shows no evidence for this hypothesis, with Web use generally associated with increased health professional contact (though this may also raise concerns regarding unnecessary utilisation of health resources) (Andreassen et al., 2007; Couper et al., 2010; C. J. Lee, 2008; Suziedelyte, 2012; Valero-Aguilera et al., 2012).

Despite the many advantages and the widespread use of the Web as a health information source, this is associated with several problems. The main difficulties encountered by users are discerning trustworthy from less trustworthy information (Dickerson, Reinhart, Boehmke,
Akhu-Zaheya, 2011; Sommerhalder, Abraham, Zufferey, Barth, & Abel, 2009), feeling overwhelmed by the large amount of information (Chen, 2012; Sommerhalder et al., 2009) and difficulty accessing relevant information, particularly among older users and those with low health literacy (Agree, King, Castro, Wiley, & Borzekowski, 2015; Manganello et al., 2016).

Because the Web has the potential to be a valuable and convenient source of health-related information but is associated with various problems and difficulties, the role of the Web in people’s decision-making regarding health issues is unclear. Furthermore, to my knowledge, there is to date no study examining Web use prior to diagnosis among lung cancer patients, and the potential impacts of Web use on help-seeking.

1.5 In summary

To summarise, survival rates among lung cancer patients are low, and this has been attributed, in part, to prolonged intervals between first symptom appearance and presentation of patients to health services. Strategies are needed to help those afflicted with lung cancer evaluate their symptoms and seek help promptly. Web-based information could play a key role in achieving this, as many people research their symptoms using Internet search engines. The Web shows potential in acting as a health information source, but it involves many difficulties and limitations. Understanding how individuals use the Web to appraise their lung cancer symptoms could help us to utilise this information source more effectively, thus reducing delay and positively impacting on survival chances.

1.6 Thesis aim

In this thesis, I aimed to explore the role of the Web in help-seeking for lung cancer symptoms. I aimed to investigate whether the Web can play a role in the time before diagnosis of lung cancer, and whether it can be used to disseminate useful information to those with relevant symptoms, in order to facilitate early presentation to health services. To address this aim, I pursued the following objectives:

- To establish whether people with lung cancer access the Web to appraise their symptoms before diagnosis
• To explore whether there is potential to harness the Web as a health information source to promote earlier help-seeking

• To identify strategies for presenting online health information that can facilitate earlier help-seeking, which can inform the development of health websites and online applications.

1.7 Thesis structure

This thesis comprises eight chapters and four empirical studies.

Having provided background information and clarified the research question and aims in the present chapter, I systematically review relevant literature in Chapter 2 in order to demonstrate the novelty of the research and show how previous research informed my subsequent work. The systematic review critically appraises methods used in previous research to examine Web use for symptom appraisal, as well as synthesising main findings, taking study limitations into account.

Chapter 2 concludes with a discussion of how findings from the systematic review relate to the context of lung cancer, which then leads to the first empirical study in this thesis. Thus Chapter 3 presents findings from an exploratory mixed-methods study conducted with recently diagnosed lung cancer patients and proxies.

In Chapter 4, I describe how I used insights from Chapter 2 and 3, as well as previous research, psychological theory, and patient and public involvement work, to develop a Web-based intervention to promote help-seeking for lung cancer symptoms. Chapter 5 shows how I tested this intervention in a feasibility study which aimed to assess whether the proposed study design, the intervention website, and recruitment strategies were acceptable and feasible, and to obtain estimates for sample size calculations.

After completion of the feasibility study, the intervention was evaluated in an online randomised controlled trial (RCT), which is described in Chapter 6. Because this study concluded with the identification of further limitations, I conducted a further online RCT, presented in Chapter 7.

Finally, Chapter 8 brings together findings from the systematic review and the four empirical studies and critically appraises their strengths and limitations, highlighting main findings and implications for future research and practice.
Chapter 2: Web use for symptom appraisal for physical health conditions: a systematic review

2.1 Introduction

The material presented in this chapter has been published in (Mueller, Jay, Harper, Davies, et al., 2017), but is partly rewritten and expanded upon for this thesis.

As the previous chapter highlighted, there is potential for the Web to act as a health information source prior to a diagnosis of lung cancer. Before this can be effectively utilised, however, we need to gain an understanding of the role of this information source prior to diagnosis, how individuals use the Web, and what consequences this can entail. In this chapter, I therefore present the findings drawn from a systematic review of the literature on Web use prior to diagnosis.

Previous research on Web use among cancer populations has mainly focused on post-diagnosis searches (Bylund et al., 2012; Chou et al., 2011; Dickerson et al., 2011; Finney Rutten et al., 2016; Jabbour et al., 2017; Kirschning & von Kardorff, 2008; C. Lee, Ramirez, Lewis, Gray, & Hornik, 2012; Mattsson, Olsson, Johansson, & Carlsson, 2017; McLeod, Yu, & Ingledew, 2017; Peterson & Fretz, 2003; Shaverdian et al., 2017; Valero-Aguilera et al., 2012).

However, the way the Web is used by patients who have obtained a specific diagnosis from a healthcare professional is likely to differ from the way it is used in the absence of a professional diagnosis, when appraising symptoms. Post-diagnosis, individuals have specific medical terms they can use as search terms. Most focus their Web search on treatment options, illness management and prognosis (Dickerson et al., 2011; Kirschning & von Kardorff, 2008). When appraising symptoms with the aim of diagnosing them, on the other hand, most individuals have only symptoms and lay medical knowledge to guide their search, and symptoms are sometimes vague and difficult to describe.

Despite these important distinctions, pre-diagnosis Web searches remain under-researched. For this review it was therefore necessary to broaden the scope to other disease types rather than focusing solely on cancer, and examine Web use for symptom appraisal for any physical health condition.
One of the aims of this PhD project is to explore whether the Web can be utilised to facilitate help-seeking for lung cancer symptoms. It is unclear whether Web use for symptom appraisal generally plays a beneficial or detrimental role in healthcare delivery. For example, Web use for symptom appraisal may cause anxiety about health by making individuals falsely believe they have a serious condition when they do not. On the other hand, it may encourage people with warning signs to present to health services, promoting earlier diagnosis. It could also falsely reassure people that symptoms are not serious, thus preventing earlier diagnosis. This is particularly important for potentially life-threatening or debilitating conditions which are easier to treat when detected early, such as lung cancer (Radzikowska et al., 2012).

To utilise this potential for reducing strain on healthcare resources and promoting earlier diagnosis, we need to understand the current role of the Web in help-seeking and symptom evaluation, and the strategies people use to access information, taking differing contexts into account. As these questions cannot be addressed in a single study, a systematic review is required, involving a thorough and comprehensive search of the literature, critical appraisal of individual studies, and extraction and synthesis of relevant findings. In the following, I present the findings of this systematic review, and I go on to discuss the relevance of the findings in the context of the present PhD project on lung cancer.

This systematic review addresses the following five review questions:

RQ1: What proportion of different populations (e.g. general, specific disease or demographic groups etc.) use the Web to appraise symptoms?

RQ2: Which symptoms are likely to be researched online?

RQ3: How is Web use for symptom appraisal conducted (search strategies)?

RQ4: What are the behavioural consequences of Web use for symptom appraisal?

RQ5: What are the emotional consequences of Web use for symptom appraisal?

2.2 Methods

A protocol was developed within the research team based on the review questions and an initial broad search of the available literature, using the Centre for Reviews and Dissemination’s
guidance for undertaking reviews in healthcare (Centre for Reviews and Dissemination, 2009) and the PRISMA statement for reporting systematic reviews (Liberati et al., 2009).

2.2.1 Eligibility criteria

Study focus

I included studies that addressed use of the Web to appraise symptoms, i.e. to research symptoms and their potential causes. This could include both actual symptoms and symptoms in fictional scenarios. This did not have to be the primary focus of the study; some reference to Web use for symptom appraisal was sufficient. If studies examined health-related Web use in general, they were screened during full-text review and excluded if no specific reference to symptom appraisal was made. Studies that analysed anonymous logs were included if they examined symptom-related searches. I included only studies that focused on human behaviour; studies which evaluated the performance of Web-based tools were excluded.

Populations

Studies on Web use for symptom appraisal of any physical health conditions were included. Studies examining mental health/psychiatric conditions were excluded in order to focus the scope of the review on physical health conditions. Studies on Web use by health professionals were excluded. Studies from any country were included, as long as the publication was written in English.

Study design

The initial scoping suggested a scarcity of research in this area, thus I did not limit included studies to any particular design. Non-empirical studies, e.g. theoretical papers and literature reviews, were excluded.

Publication types

Full paper, English language publications were included, regardless of the original language of the research.
2.2.2 Information sources

Pubmed, Embase, PsycInfo, ACM digital library, SCOPUS, and Web of Science were searched for relevant publications up to October 2017. To minimise publication bias, grey literature was explored by searching OpenGrey, an open-access database containing over 700,000 bibliographical references of grey literature, as well as Google Scholar. I also searched the British Library Integrated Catalogue, which contains reports, conference abstracts and theses. Finally, authors in the field were contacted to inquire about any unpublished material, if two or more of their papers were among the included studies, or if their paper was judged as particularly relevant to the review (e.g. if examining Web use for symptom appraisal was the primary focus of the study).

2.2.3 Search

The terms Internet, Web, online, search engine, Google, help-seeking, health information-seeking, symptom*, and diagnosis were entered into the databases (note Google was used as a search term as this is by far the most widely used search engine worldwide (eBiz MBA, n.d.)).

The search strategy is shown in Appendix 1.

2.2.4 Study selection

The study selection process followed the guidelines provided in the PRISMA statement (Liberati et al., 2009). Search terms were entered into the databases, and all returned studies were imported into a single Mendeley file. Three independent reviewers (JM, AD and JV) assessed the studies for eligibility, such that each study was assessed by at least two reviewers. Studies were first screened by titles and abstracts. Selected studies were then screened for inclusion by reading full texts. Any discrepancies between the reviewers were discussed until consensus was reached. Reference lists of included studies were hand-searched for further eligible studies. I also hand-searched journals if they contained ≥ 2 papers included in this review, or if the general journal topic area was particularly relevant to the review, to ensure inclusion of studies not yet loaded on electronic databases.

1 Recent publications have been included to render the systematic review chapter complete and up-to-date, but it should be noted that publications later than 2015 did not feed into intervention development. Chapter 4 clearly outlines which findings from the systematic review were used to inform intervention development.
2.2.5 Data collection process and data items

From each study, any information regarding use of the Web for symptom appraisal was extracted, as well as details on study design, procedure, population, sampling method, entry and inclusion criteria for study participants, sample size, measures, and details of analysis methods (data extraction sheet in Appendix 2).

2.2.6 Quality appraisal

A quality appraisal of selected papers was conducted based on five criteria designed for reviews incorporating mixed study designs (Dixon-Woods et al., 2006) (Appendix 3). Quality appraisal involved two stages. First, papers were assessed for inclusion in the review using a relatively liberal threshold; papers were scored eligible if they addressed each criterion at least minimally. Criteria were then applied more rigorously using a 3-point scoring system (low/medium/high – see Appendix 3), and main limitations of each study were identified. This assessment was used to critically appraise studies during synthesis of the findings.

2.2.7 Synthesis of results

The extracted data were synthesised using Thematic Analysis, one of the main approaches used to review and synthesise qualitative and quantitative evidence (Dixon-Woods, Agarwal, Jones, Young, & Sutton, 2005; Mays, Pope, & Popay, 2005).

My analysis involved the following steps (Braun & Clarke, 2008):

1. Data familiarisation: Familiarisation with the data was achieved by reading all included studies several times and extracting the relevant information into the data extraction sheets.

2. A priori grouping: Data from the data extraction sheets were grouped according to the review question they pertained to, and summarised in a matrix. Studies were entered into the rows of the matrix, while study characteristics, limitations, and review questions were entered into the columns. This matrix enabled us to compare the findings of different studies pertaining to the same review question, taking methodological aspects into account (example matrix shown in Appendix 4). This method was adapted from Framework Analysis, which is a specific form of Thematic Analysis (Spencer, Ritchie, O’Connor, Morrell, & Ormston, 2013).
3. Generation of initial codes: The data were initially coded using semantic codes within the NVivo10 environment (Version 10), using the matrix to compare results across studies.

4. Searching for themes: Once all data extracts were coded, codes were sorted into broader, more conceptual categories to create themes.

5. Reviewing themes: Finally, I reviewed the data extracts the themes relate to, to determine whether the created themes satisfactorily capture the raw data.

For quantitative studies reporting proportions without confidence intervals, 95% confidence intervals were computed using the Wilson method (Brown, Cai, & Dasgupta, 2001) to facilitate comparisons between studies.

2.3 Results

2.3.1 Study selection

Thirty-four\(^2\) studies were identified as eligible for inclusion in the review. The search process is illustrated in Figure 2.1. Hand-searching reference lists led to 6 further inclusions, and the grey literature search yielded one further inclusion (a dissertation which had not been published in form of articles). The Journal of Medical Internet Research, the Journal of Health Communication, Telemedicine and e-Health, and the Journal of the American Medical Informatics Association were hand-searched, resulting in 15 full text assessments, and two further inclusions. Four authors were contacted to enquire about unpublished material. I received one reply, concerning a paper I had already identified.

\(^2\) Two publications based on the present PhD would be eligible for inclusion in the systematic review (Mueller, Davies, Harper, Jay, & Todd, 2016; Mueller, Jay, Harper, & Todd, 2017); however this would create a circular referencing issue because subsequent chapters describe how findings from the systematic review were used to inform the development of future studies in this PhD. Therefore, these two publications are not included in the systematic review for the purpose of this thesis, but this work will be discussed in detail in subsequent chapters.
2.3.2 Study characteristics

Table 2.1 provides an overview of study characteristics. Study designs included cross-sectional surveys (n=13, one of these with an embedded qualitative interview study, one with an embedded observational study), qualitative interview and focus group studies (n=5), experimental studies (n=6), studies evaluating questions posed by users to health website (n=2), a study evaluating clicks on a specific health website (n=1) and studies involving the analysis of log data from search engines (n=7). Two of these combined a log-based approach with a survey. Most studies were conducted in the USA (n=24).
As Table 2.2 shows, some studies explored Web use regarding current symptoms (n=5) or symptoms that had been experienced previously (n=7), while other studies examined Web use for symptom appraisal by providing participants with a symptom vignette, and instructing them to imagine they have these symptoms (n=9). In several studies (n=11), the exact situation of participants was unclear, because anonymous data were collected online, and in two cases the sample was mixed. Table 2.2 also highlights the variety and non-specificity of symptoms examined; most studies (n=16) examined general symptoms, and while 10 studies examined specific conditions, only three studies examined similar conditions (Hay et al., 2008; Powley et al., 2016; Simons et al., 2017). Finally, most studies (n=21) did not follow up whether participants had received a diagnosis.

### 2.3.3 Quality assessment and risk of bias

Quality assessment of the studies is shown in Appendix 3. Below I use this information to critically appraise evidence from the studies and assess risk of bias.
<table>
<thead>
<tr>
<th>Author, date</th>
<th>Study design</th>
<th>Aim</th>
<th>Study population</th>
<th>Setting</th>
<th>Sample size</th>
<th>Nature of measures &amp; Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attfield, Adams &amp; Blandford (2006)</td>
<td>Qualitative interview study</td>
<td>Cross sectional Explore information seeking of patients before and after consultations, its situational influences and its impact on patient-provider relationships.</td>
<td>2 groups of 8 NHS patients: One group from a Patient Advice and Liaison Service (PALS) patient panel (43-81 years, average 64) and one group of MSc students for HCI (25-42 years, average 31)</td>
<td>UK</td>
<td>16</td>
<td>Semi-structured interviews, eliciting accounts of health information seeking episodes and how they relate to on-going healthcare.</td>
</tr>
<tr>
<td>Briet et al. (2014)</td>
<td>Quantitative cross sectional analysis of website queries</td>
<td>Explore the nature and content of questions and answers on a health website, and to examine the situations of patients asking questions.</td>
<td>Users asking hand surgery-related questions from a free online health consultation website</td>
<td>USA (American website – no restriction regarding location of website users)</td>
<td>131 questions</td>
<td>Questions and answers to a health website were categorized and analysed descriptively.</td>
</tr>
<tr>
<td>Cartright, White &amp; Horvitz (2011)</td>
<td>Cross sectional log based study</td>
<td>Analyse the search activity of users researching health information online, and identify goals and patterns of search behaviour.</td>
<td>A set of filtered logs from a toolbar deployed by the Windows Live search engine, containing at least one symptom.</td>
<td>USA (English language logs, but no restriction regarding location of users)</td>
<td>2,329,231 actions (=queries issued to a search engine)</td>
<td>Logs were mined and categorized as either evidence-directed, hypothesis-directed with diagnostic intent, or hypothesis-directed with informational intent, according to defined algorithms.</td>
</tr>
<tr>
<td>Author, date</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population</td>
<td>Setting</td>
<td>Sample size</td>
<td>Nature of measures &amp; Procedure</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
<td>-----</td>
<td>------------------</td>
<td>---------</td>
<td>-------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>Chin (2009)</td>
<td>Experimental between subjects design: 2x2 (ill-well defined tasks, younger-older users)</td>
<td>Compare older and younger adults in their performance and search behaviour in ill and well defined tasks.</td>
<td>Younger and older adults from a University community</td>
<td>USA</td>
<td>69; 41 younger adults (18-35), 28 older adults (60-83)</td>
<td>Participants were randomized to complete either an ill-defined task (find possible causes for a list of symptoms) or well-defined task (find a specific medical term), using a health website. Cognitive measures (working memory capacity, processing speed); Health literacy measures, Medical knowledge measure, search performance for both tasks were measured.</td>
</tr>
<tr>
<td>Chin &amp; Fu (2010)</td>
<td>Experimental between subjects design: 2x2x2 (older-younger adults, parts-systems interface, parts-system task)</td>
<td>Examine differences between older and younger adults in interacting with different online search tasks and interfaces.</td>
<td>Younger and older adults from community of a medium sized city</td>
<td>USA</td>
<td>46, 23 younger (18-28) and 23 older (60-77) adults</td>
<td>Participants were given a symptom vignette and asked to find possible causes. Participants were randomized to complete either a Parts task (described symptoms based on body parts) or a Systems task (described symptoms by functional systems). Tasks were completed either in the Parts interface (categorised symptoms based on body parts) or Systems interface (categorised symptoms based on functional body systems). Measures included Patients’ Medical Background Knowledge, Mental Interface Match Index, Broadness (no. of links), link decision time: time spent reading.</td>
</tr>
<tr>
<td>Author, date</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population</td>
<td>Setting</td>
<td>Sample size</td>
<td>Nature of measures &amp; Procedure</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Chiu (2016)</td>
<td>Analysis of inquiries posted to a health website</td>
<td>Investigate questions posted to a medical consultation website</td>
<td>Users posting questions to a medical consultation website</td>
<td>Taiwan (Taiwanese website but no restrictions on locale of users)</td>
<td>1200 randomly selected consultation Q&amp;As</td>
<td>A subset of questions to the consultation website were randomly selected and analysed in terms of length, moment of seeking consultation, communication strategies, purpose and identity disclosure.</td>
</tr>
<tr>
<td>Cooper, Polonec, Steward &amp; Gelb (2013)</td>
<td>Qualitative study (focus groups)</td>
<td>Explore how women would evaluate symptoms associated with gynaecologic cancers.</td>
<td>Women aged 40–60 years</td>
<td>USA</td>
<td>132</td>
<td>Discussion in focus groups: Which symptoms from a list would be of most concern, why, and what could cause them, what would be their hypothetical response to them, what were actual responses in the past.</td>
</tr>
<tr>
<td>Cumming et al. (2010)</td>
<td>Cross-sectional Web-based survey study</td>
<td>Evaluate digital storytelling videos (videos of people talking about their own experiences) about help-seeking for menopausal symptoms.</td>
<td>Visitors of a UK-based menopause website</td>
<td>UK (UK website - no restriction regarding location of website users)</td>
<td>539</td>
<td>Participants viewed a storytelling video online, and then completed a questionnaire evaluating the effect of the video on feeling informed, planned future help-seeking etc.</td>
</tr>
<tr>
<td>De Choudhury, Morris &amp; White (2014)</td>
<td>Cross-sectional survey study (quantitative + qualitative data) +longitudinal log-based study</td>
<td>Research the prevalence of health activities on social media and search engines; characterize health activities on the different platforms and describe how people evaluate information obtained from these.</td>
<td>Survey: US adults 18-70 (census representative sampling)</td>
<td>USA</td>
<td>210 survey respondents; 125,166,549 tweets; 174,605,024 searches</td>
<td>In the survey, participants were asked questions about their experiences using Twitter and search engines to share and seek health information. In the log analysis, Tweets and logs were categorized as relating to 4 categories: (1) symptoms of major diseases, (2) benign explanations (non-life-threatening illnesses), (3) serious illnesses, and (4) disabilities. Logs were then analysed descriptively.</td>
</tr>
<tr>
<td>Author, date</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population</td>
<td>Setting</td>
<td>Sample size</td>
<td>Nature of measures &amp; Procedure</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>-----</td>
<td>------------------</td>
<td>---------</td>
<td>-------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>Fiksdal et al. (2014)</td>
<td>Qualitative focus group study</td>
<td>To gain a deeper understanding of online health-searching behavior in order to inform future developments of personalizing information searching and content delivery.</td>
<td>Adult, English-speaking members of the Olmsted County, MN community (where Mayo Clinic is located) and Mayo Clinic patients, employees, and family visitors.</td>
<td>USA</td>
<td>19</td>
<td>Moderators used a semi-structured moderator guide to facilitate discussion in focus groups about: (1) participants’ perception and understanding of health care information, (2) the process of information collection on the Internet, (3) understanding and usage of information, and (4) implications of health care information for health and well-being.</td>
</tr>
<tr>
<td>Fox &amp; Duggan (2013)</td>
<td>Nationwide cross sectional survey</td>
<td>The Pew Internet &amp; American Life Project is an initiative of the Pew Research Center, a nonprofit “fact tank” that provides information on the issues, attitudes and trends shaping America and the world.</td>
<td>Adults living in the United States</td>
<td>USA</td>
<td>3,014</td>
<td>People were contacted via telephone for telephone interviews about online health information seeking.</td>
</tr>
<tr>
<td>Hay et al. (2008)</td>
<td>Mixed-methods survey and interview study</td>
<td>Understand the extent and reasons for online research prior to first appointments for patients in a rheumatology clinic</td>
<td>English speaking US adults (≥17 years)</td>
<td>USA</td>
<td>120</td>
<td>Before their appointment, patients were interviewed about online health information (OHI) seeking, and completed the Wong-Baker-Faces Pain Scale. The consultation was audio-recorded to determine whether OHI was mentioned. Afterwards, patients completed a satisfaction scale regarding the consultation.</td>
</tr>
<tr>
<td>Author, date</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population</td>
<td>Setting</td>
<td>Sample size</td>
<td>Nature of measures &amp; Procedure</td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------</td>
<td>---------</td>
<td>-------------</td>
<td>--------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Keselman, Browne &amp; Kaufman (2008)</td>
<td>Cross-sectional qualitative interview and think aloud study.</td>
<td>Explore users’ information seeking difficulties by conceptualizing information seeking as a form of hypothesis testing, and to examine the role of users competencies in online information seeking.</td>
<td>Lay individuals (convenience sample)</td>
<td>USA</td>
<td>20</td>
<td>Participants read a hypothetical scenario describing a relative who experienced symptoms typical of stable angina. Then discussion of possible causes of symptoms from the symptom vignettes in semi-structured interviews; then Think Aloud while they researched symptoms on MedlinePlus.</td>
</tr>
<tr>
<td>Lauckner &amp; Hsieh (2013)</td>
<td>Experimental 2x2 design (position top – bottom, frequency high – low)</td>
<td>Does the position and frequency of serious conditions in search results affect perceived severity and susceptibility, and is this related to negative emotional outcomes? Do health literacy and experience with online health seeking moderate these relationships?</td>
<td>Students from an undergraduate communication course at a large Midwestern university</td>
<td>USA</td>
<td>274</td>
<td>The study took place online. Participants were presented with a symptom vignette. They were then presented with a search engine result page, manipulated to show serious conditions either at the top or bottom, and low or high frequency of serious conditions. Participants then completed several scales: Perceptions of severity and susceptibility using the Risk Behavior Diagnosis scale; History of viewing online health information; their health status, how often they experienced each of the four symptoms; and their demographic information. Health literacy using the Newest Vital Sign (NVS).</td>
</tr>
<tr>
<td>Author, date</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population</td>
<td>Setting</td>
<td>Sample size</td>
<td>Nature of measures &amp; Procedure</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------</td>
<td>-------------</td>
<td>-------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Luger (2014)</td>
<td>Experimental 2x2 design: two different symptom vignettes (mononucleosis or scarlet fever), either Google or WebMD</td>
<td>Explore older adults’ online health seeking to determine the cognitive and diagnostic processes involved.</td>
<td>Older US adults, 50 years or older, community resident, without cognitive impairment, who owned a computer</td>
<td>USA</td>
<td>79</td>
<td>Participants were presented with one of two symptom vignettes, and asked to diagnose them using Think Aloud, either on Google or WebMD. Measures taken included: Screen Shots, Illness Perception Questionnaire, Concurrent Think-Aloud, Interactivity Scale (perceived interactivity of websites), Perceived Cognitive Effort, Perceived Competence and Choice. Computer Experience self-report questionnaire. Recent and Lifetime Health History. BDI. Neuroticism.scale from Big 5 Inventory. Preferred Role in Medical Encounters.</td>
</tr>
<tr>
<td>Medlock et al. (2015)</td>
<td>Cross-sectional online survey</td>
<td>To determine which information resources seniors who use the Internet use and trust for health information, which sources are preferred, and which sources are used by seniors for different information needs.</td>
<td>Members of a local senior (Christian) organisation</td>
<td>Netherlands</td>
<td>118</td>
<td>Participants completed an online questionnaire, which included questions about health information resources used. The Autonomy Preference Index was used to assess information needs and preferences for involvement in health decisions.</td>
</tr>
<tr>
<td>Morgan et al. (2014)</td>
<td>Analysis of inquiries posted to a health website</td>
<td>Describe what information people seek from a US website about genetic and rare diseases, and why.</td>
<td>Random sample of English-language inquiries posted by lay people to the question and answer section of the GARD website and inquiries sent via email.</td>
<td>USA (American website but no restrictions on locale of users)</td>
<td>278 inquiries, 68 from 2006 and 210 from 2011</td>
<td>A random sample of questions posted to the GARD website were analysed thematically. Collected data included: Inquiry origin (domestic), type of contact (email and Web-based form), gender, date received at the information centre the specific condition for which they were inquiring, primary language (English), and their reason for inquiry.</td>
</tr>
<tr>
<td>Author, date</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population</td>
<td>Setting</td>
<td>Sample size</td>
<td>Nature of measures &amp; Procedure</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
<td>-----</td>
<td>------------------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Norr et al., 2014 (2014)</td>
<td>Experimental within-subjects design</td>
<td>Investigate whether viewing medical websites adversely affects anxiety sensitivity (AS).</td>
<td>Undergraduate students from a large university in the Southern United States.</td>
<td>USA</td>
<td>56</td>
<td>Participants first completed the Anxiety Sensitivity Index (ASI), Intolerance of Uncertainty Scale (IUS), and a health anxiety scale (SHAI). Participants were randomized to view either symptom-related websites or general health and wellness control websites. Afterwards, they completed the ASI and SHAI.</td>
</tr>
<tr>
<td>North et al. (2011)</td>
<td>Cross-sectional analysis of clicks on a health website and calls to a telephone triage system.</td>
<td>Establish what symptoms internet users tend to look up online, and whether telephone triage algorithms could be applied to these.</td>
<td>All symptom assessment callers to Ask Mayo Clinic (telephone triage) and all clicks to specific symptoms on the symptom checker page of MayoClinic.com</td>
<td>USA</td>
<td>70 370 calls; 2 059 299 clicks.</td>
<td>For the MayoClinic website, click data was collected using Google Analytics; for the telephone triage, all completed calls were counted and put into symptom categories based on the algorithm/guideline used during the call.</td>
</tr>
<tr>
<td>Perez et al. (2015)</td>
<td>Experimental study with Think Aloud</td>
<td>Describe Internet search processes and identify demographic and personal characteristics associated with use of System 1 (does not include hypothesis testing and evidence gathering) and System 2 (includes hypothesis testing and evidence gathering) processing.</td>
<td>Young adults aged 21-35 with experience of online health information and reported barriers to accessing health services.</td>
<td>USA</td>
<td>78</td>
<td>Participants were randomised to one of two symptom scenarios and instructed to search the Internet while using Think Aloud. Participants’ Internet searches and think-out-loud vocalizations were digitally recorded using screen capture video-recording software.</td>
</tr>
<tr>
<td>Author, date</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population</td>
<td>Setting</td>
<td>Sample size</td>
<td>Nature of measures &amp; Procedure</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
<td>-----</td>
<td>------------------</td>
<td>---------</td>
<td>-------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>Powell et al. (2011)</td>
<td>Cross-sectional survey with embedded qualitative semi-structured interviews.</td>
<td>Identify the characteristics and motivations of online health information seekers accessing the NHS Direct website</td>
<td>Users of the NHS Direct website</td>
<td>UK</td>
<td>792 for survey, 26 for interviews</td>
<td>Users of the NHS Direct website completed an online questionnaire survey. A subsample of survey respondents participated in in-depth, semistructured, qualitative interviews by telephone or instant messaging/email.</td>
</tr>
<tr>
<td>Powley et al. (2016)</td>
<td>Cross-sectional survey and observational study</td>
<td>Evaluate whether patients with inflammatory arthritis and inflammatory arthralgia use the internet for symptom appraisal and to assess the advice given and diagnoses suggested by the NHS and WebMD symptom checkers.</td>
<td>Newly presenting patients with either clinically apparent synovitis or a new onset of symptoms consistent with inflammatory arthritis but without clinically apparent synovial swelling attending a secondary care based rheumatology clinic.</td>
<td>UK</td>
<td>34</td>
<td>Patients completed a brief survey on internet use for symptom appraisal prior to attending clinic; patients were then asked to complete the NHS and WebMD symptom checkers based on their symptoms and their answers and the outcomes were recorded; demographic and disease-related data were obtained from clinic records.</td>
</tr>
<tr>
<td>Rice (Rice, 2006)</td>
<td>Cross-sectional survey study; secondary analysis of existing dataset</td>
<td>Understand what influences online health seeking, what the reported benefits of online health seeking are, and to identify similarities among online activities.</td>
<td>US adults: Respondents from studies conducted within the Pew Internet and American Life project.</td>
<td>USA</td>
<td>13 978 respondents in 2000 who reported health seeking online, 500 of these were telephone interviewed in 2001</td>
<td>Respondents were contacted via telephone for telephone interviews asking about online health seeking.</td>
</tr>
<tr>
<td>Simons et al. (Simons et al., 2017)</td>
<td>Qualitative interview study</td>
<td>To explore behaviours people might undertake in response to rheumatoid arthritis symptoms, compared to bowel cancer/angina</td>
<td>Patients from two inner-city UK GP practices aged 18-40, 41-60, and 61+ without a diagnosis of inflammatory arthritis</td>
<td>UK</td>
<td>32</td>
<td>Patients were presented with 4 symptom vignettes (1. Inflammatory joint symptoms, 2. Early rheumatoid arthritis symptoms, 3. Bowel cancer symptoms, 4. Angina symptoms) and asked what they would do in response.</td>
</tr>
<tr>
<td>Author, date</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population</td>
<td>Setting</td>
<td>Sample size</td>
<td>Nature of measures &amp; Procedure</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>----------</td>
<td>-------------</td>
<td>------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Teriaky, Tangi &amp; Chande (Teriaky, Tangri, &amp; Chande, 2015)</td>
<td>Cross-sectional survey</td>
<td>Understand how outpatients awaiting initial gastroenterology consultation seek medical information on the Internet and how wait times affect Internet usage.</td>
<td>Patients awaiting appointments at a general gastroenterology clinic in London, Ontario, Canada.</td>
<td>Canada</td>
<td>87</td>
<td>Patients awaiting gastroenterology consultation were asked to complete a questionnaire consisting of 16 multiple-choice questions to understand patient use of Web resources for medical information. Abstracted information included patient demographics, level of education, reason for referral, preceding investigations, patient resources utilized, websites browsed, information obtained, reasons for seeking information on the Internet, patient self-diagnosis, and lifestyle changes instituted.</td>
</tr>
<tr>
<td>Thomson, Siminoff &amp; Longo (Thomson, Siminoff, Longo, M.D., &amp; L.A., 2012)</td>
<td>Cross sectional survey study</td>
<td>Explore characteristics of colorectal cancer patients who used the Web to appraise symptoms prior to diagnosis.</td>
<td>Newly diagnosed colorectal cancer patients (&lt;6 months)</td>
<td>USA</td>
<td>242</td>
<td>Semi-structured interviews focused on patient sociodemographic and psychological factors, symptom recognition and appraisal, and communication with healthcare providers, friends, and family.</td>
</tr>
<tr>
<td>Author, date</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population</td>
<td>Setting</td>
<td>Sample size</td>
<td>Nature of measures &amp; Procedure</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td>--------------------------</td>
<td>----------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>White &amp; Horvitz (White &amp; Horvitz, 2009a)</td>
<td>Longitudinal log-based study and cross-sectional survey.</td>
<td>1) Describe escalations that occur when users search for common symptoms and this escalates to queries about serious conditions, and 2) examine how this persists over several sessions.</td>
<td>Log data related to symptom queries (no mention of restriction by locale) from all major Web search engines such as Google, Yahoo!, or Live Search: randomly selected employees of the Microsoft Corporation who had performed at least 1 health-related online search. Survey: Microsoft employees</td>
<td>USA</td>
<td>Logs: 8,732 users with symptom related queries Survey: 515 participants</td>
<td>Analysis of logs: Formulated a list of symptoms and associated benign and serious conditions. Recorded all queries to search engines and clicks on result pages, and identified those that included symptoms as search terms. Escalations: Observed increases in medical severity of search terms within a search session. Non-escalations: Search progresses to benign explanation of the symptom. Survey: Microsoft employees were sent a survey with open and closed-ended questions regarding participants’ medical history and online search behaviour.</td>
</tr>
<tr>
<td>White &amp; Horvitz (White &amp; Horvitz, 2009b)</td>
<td>Cross-sectional survey study.</td>
<td>Explore how lay individuals use the Web to find explanations for symptoms, what activities they pursue and what their experiences are.</td>
<td>5000 Microsoft employees were invited via email, from these volunteers were chosen who indicated in a pre-screening that they searched the Web for medical information</td>
<td>USA</td>
<td>515 survey respondents</td>
<td>Microsoft employees were sent a survey to elicit perceptions of online medical information, experiences in searching for this information, and the influence of the Web on healthcare concerns and interests. The survey contained ‘around 70’ open and closed questions.</td>
</tr>
<tr>
<td>Author, date</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population</td>
<td>Setting</td>
<td>Sample size</td>
<td>Nature of measures &amp; Procedure</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------</td>
<td>-----------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>White &amp; Horvitz</td>
<td>Cross sectional log study using logs</td>
<td>Predict escalations in searches based on characteristics of websites</td>
<td>Logs from windows live browser toolbar, English speaking USA relating to 6 basic symptoms.</td>
<td>USA (log data issued from US locale)</td>
<td>“Many thousands of logs were mined”</td>
<td>Cases were identified where queries for symptoms were followed by a query about a related serious condition. Cases where it led to a benign query or no change were termed non-escalations. Using logistic regression, a model was developed to predict escalation using website features of the previously visited page. Website features: Structural features, title &amp; URL features, 1st person testimonials, page reliability/credibility, commercial intent…</td>
</tr>
<tr>
<td>(White &amp; Horvitz, 2010a)</td>
<td>from windows live toolbar.</td>
<td>visited.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White &amp; Horvitz</td>
<td>Longitudinal log based study.</td>
<td>Establish predictors of when searches for common symptoms lead to health care utilization.</td>
<td>Logs from consenting Windows live toolbar users over a 6 month period relating to 3 symptoms: chest pain, muscle twitches, abdominal pain</td>
<td>USA (log data issued from US locale)</td>
<td>700 queries with symptom HUI transition, 700 queries with symptoms no HUI transition.</td>
<td>Log analysis: Logs containing symptoms as search terms were filtered, and it was determined whether subsequent searches showed Healthcare Utilisation Intent (HUI). Logistic regression was used to predict HUI based on search characteristics. Log entries include a user identifier, a timestamp for each page view, and the URL of the page visited. HUI: queries that indicate searching for contact information for medical facilities</td>
</tr>
<tr>
<td>(White &amp; Horvitz, 2010b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, date</td>
<td>Study design</td>
<td>Aim</td>
<td>Study population</td>
<td>Setting</td>
<td>Sample size</td>
<td>Nature of measures &amp; Procedure</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------</td>
<td>---------</td>
<td>-------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>White &amp; Horvitz (White &amp; Horvitz, 2012)</td>
<td>Longitudinal log-based study</td>
<td>Explore how users search for medical concerns and particularly how these concerns impact on future behaviour, e.g. how this influences focus and attention of future searches.</td>
<td>Logs from consenting Windows live toolbar users over a three month period.</td>
<td>USA</td>
<td>169,513 queries</td>
<td>Queries were labeled to identify medical and symptoms related queries, and escalations. Subsequently occurring searches were examined. Log entries included a unique user identifier, a timestamp for each page view. Search sessions on Google, Yahoo!, and Bing. Escalation queries were categorized as within-session and between session.</td>
</tr>
<tr>
<td>White &amp; Horvitz (White &amp; Horvitz, 2013a)</td>
<td>Longitudinal log-based study</td>
<td>Research questions: 1) whether snippets in search results are biased towards serious conditions when symptoms are entered into search engines and 2) how these snippets influence user behaviour.</td>
<td>Log data related to symptoms queries generated in English speaking US locale.</td>
<td>USA</td>
<td>2,070 symptom queries from 714 users</td>
<td>Log data relating to symptom queries were filtered. Subsequent behavior on the search engine result page was examined, including hovering, cursor movements, clicks, scrolling, as well as bounding boxes of areas of interest (AOIs).</td>
</tr>
<tr>
<td>Yastik (Yastik, 2017)</td>
<td>Cross-sectional survey</td>
<td>Determine demographics and symptoms of non-emergent Emergency Department (ED) patients who engage in Internet health information gathering, describe information gathering, and explore whether it influences presentation to ED</td>
<td>Non-emergent ED patients at a US hospital who engaged in Web use related to their presenting symptom(s) prior to presenting to the ED</td>
<td>USA</td>
<td>74</td>
<td>ED patients were recruited in the ED lobby/registration area, and completed a survey assessing demographics, symptoms (type, onset, duration, perceived severity), online information gathering (timing of searches, time spent searching, sources used, discussion with others), seeking medical services (ED)</td>
</tr>
<tr>
<td>Ybarra &amp; Suman (Ybarra &amp; Suman, 2006)</td>
<td>National, longitudinal telephone survey</td>
<td>Examine which factors predict whether a Web user is likely to contact a health professional.</td>
<td>Americans living throughout the 50 states and the District of Columbia.</td>
<td>USA</td>
<td>Year 1 = 2104, Year 4: 2010, 570 of these were Year 1 participants</td>
<td>Respondents were contacted via telephone and completed a telephone survey about online health information seeking and help-seeking behaviour (seeking help from a health professional or others).</td>
</tr>
<tr>
<td>Author, date</td>
<td>Were participants symptomatic, asymptomatic, or previously symptomatic?</td>
<td>Type of symptoms examined</td>
<td>Did the study follow up whether Web use for symptom appraisal was followed by a diagnosis?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Briet et al. (2014)</td>
<td>Unclear, participants were users asking questions about symptoms (could be own symptoms, or asking on behalf of someone else)</td>
<td>Hand-illness related symptoms</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cartright, White &amp; Horvitz (2011)</td>
<td>Unclear, participants were users issuing symptom-related queries to a search engine (could be own symptoms, or asking on behalf of someone else)</td>
<td>General (any queries related to a comprehensive list of symptoms from the Merck medical dictionary)</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chin (2009)</td>
<td>Asymptomatic, participants were presented with a symptom vignette</td>
<td>Symptom vignettes included: pain and stiffness in the body; burning, itching, and sometimes tingling sensation on their body; feeling feverish and chilly after an overseas trip; fatigue, sudden weight gain and difficulty dealing with cold; however results were not analysed separately for different symptoms</td>
<td>Not applicable, as patients were not symptomatic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chin &amp; Fu (2010)</td>
<td>Asymptomatic; participants were presented with a symptom vignette</td>
<td>General (participants received 6 different vignettes with different symptoms, not assessed separately)</td>
<td>Not applicable, as patients were not symptomatic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chiu (2016)</td>
<td>Unclear, participants were users issuing queries to a health website (could be own symptoms, or on behalf of someone else)</td>
<td>General</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, date</td>
<td>Were participants symptomatic, asymptomatic, or previously symptomatic?</td>
<td>Type of symptoms examined</td>
<td>Did the study follow up whether Web use for symptom appraisal was followed by a diagnosis?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------------------------------------------------------</td>
<td>---------------------------</td>
<td>-----------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooper, Polonec, Steward &amp; Gelb (2013)</td>
<td>Asymptomatic; participants were presented with a list of symptoms</td>
<td>Symptoms related to gynaecologic cancers</td>
<td>Not applicable, as patients were not symptomatic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumming et al. (2010)</td>
<td>Most symptomatic (448/492), but some asymptomatic (44/492)</td>
<td>Menopausal symptoms</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>De Choudhury, Morris &amp; White (2014)</td>
<td>Unclear, participants were users issuing symptom-related Tweets and queries to a search engine (could be own symptoms, or on behalf of someone else)</td>
<td>General, logs were filtered for references to symptoms using a comprehensive list of symptoms from the Merck medical dictionary</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fiksdal et al. (2014)</td>
<td>Previously symptomatic</td>
<td>General (any symptoms)</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fox &amp; Duggan (2013)</td>
<td>Previously symptomatic</td>
<td>General (any symptoms)</td>
<td>Participants were asked whether their diagnosis was confirmed by a health professional; 45% said it was confirmed, 35% did not present, 19% said it was not confirmed/inconclusive</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hay et al. (2008)</td>
<td>Symptomatic; participants were newly diagnosed rheumatology patient</td>
<td>Rheumatoid symptoms</td>
<td>Yes, patients’ diagnoses were gathered after the appointment or at follow-up appointment.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keselman, Browne &amp; Kaufman (2008)</td>
<td>Asymptomatic; participants received a symptom vignette</td>
<td>Symptoms typical of stable angina</td>
<td>Not applicable, as patients were not symptomatic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lauckner &amp; Hsiehn (2013)</td>
<td>Asymptomatic; participants received a symptom vignette</td>
<td>Symptom vignettes involved one of four symptoms: headaches, chest pain, muscle twitches, or abdominal pain, BUT the different symptoms were not analysed separately</td>
<td>Not applicable, as patients were not symptomatic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luger (2014)</td>
<td>Asymptomatic; participants received a symptom vignette</td>
<td>Symptom vignettes involved either mononucleosis or scarlet fever</td>
<td>Not applicable, as patients were not symptomatic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medlock et al. (2015)</td>
<td>Previously symptomatic</td>
<td>General (any symptoms)</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, date</td>
<td>Were participants symptomatic, asymptomatic, or previously symptomatic?</td>
<td>Type of symptoms examined</td>
<td>Did the study follow up whether Web use for symptom appraisal was followed by a diagnosis?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morgan et al. (2014)</td>
<td>Unclear, participants were users issuing symptom-related Tweets and queries to a search engine (could be own symptoms, or on behalf of someone else)</td>
<td>Symptoms related to any type of genetic or rare disease</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norr et al. (2014)</td>
<td>Asymptomatic; participants viewed a list of symptom</td>
<td>General, (&quot;websites focused on symptoms of medical conditions&quot;)</td>
<td>Not applicable, as patients were not symptomatic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>North et al. (2011)</td>
<td>Unclear, participants were users searching the MayoClinic website or using a telephone triage (could be own symptoms, or on behalf of someone else)</td>
<td>General (any symptoms)</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perez et al. (2015)</td>
<td>Asymptomatic; participants received a symptom vignette</td>
<td>One of two clinical symptom scenarios: (1) fever, mild headache, dry cough, and myalgia, suggestive of influenza, and (2) fever, severe headache, and stiff neck, suggestive of meningitis.</td>
<td>Not applicable, as patients were not symptomatic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Powell et al. (2011)</td>
<td>Unclear, participants were users of the NHS website (could be own symptoms, or on behalf of someone else)</td>
<td>General (any symptoms)</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Powley et al. (2016)</td>
<td>Symptomatic; participants were patients attending a secondary care based rheumatology clinic.</td>
<td>Either clinically apparent synovitis or a new onset of symptoms consistent with inflammatory arthritis but without clinically apparent synovial swelling</td>
<td>Yes, rheumatological diagnosis was recorded after consultation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simons et al. (2017)</td>
<td>Asymptomatic; participants received a symptom vignette</td>
<td>Symptoms of rheumatoid arthritis, bowel cancer, angina</td>
<td>Not applicable, as patients were not symptomatic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teriaky, Tangi &amp; Chande (2017)</td>
<td>Symptomatic; participants were patients awaiting gastroenterology appointments</td>
<td>Symptoms related to gastroenterology</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thomson, Siminoff &amp; Longo (2012)</td>
<td>Symptomatic; participants were colorectal cancer patients</td>
<td>Symptoms related to colorectal cancer</td>
<td>Yes; all participants were diagnosed with colorectal cancer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Author, date</td>
<td>Were participants symptomatic, asymptomatic, or previously symptomatic?</td>
<td>Type of symptoms examined</td>
<td>Did the study follow up whether Web use for symptom appraisal was followed by a diagnosis?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White &amp; Horvitz (2009a)</td>
<td>Logs: Unclear, participants were users issuing symptom-related queries to a search engine (could be own symptoms, or asking on behalf of someone else)</td>
<td>Logs related to three common symptoms (headache, muscle twitches, and chest pain)</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White &amp; Horvitz (2009b)</td>
<td>Previously symptomatic</td>
<td>General (any symptoms)</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White &amp; Horvitz (2010a)</td>
<td>Unclear, participants were users issuing symptom-related queries to a search engine (could be own symptoms, or asking on behalf of someone else)</td>
<td>Queries related to any of six common symptoms: headache, chest pain, muscle twitches, abdominal pain, nausea, and dizziness</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White &amp; Horvitz (2010b)</td>
<td>Unclear, participants were users issuing symptom-related queries to a search engine (could be own symptoms, or asking on behalf of someone else)</td>
<td>Queries related to one of three symptoms: chest pain, muscle twitches, and abdominal pain.</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White &amp; Horvitz (2012)</td>
<td>Unclear, participants were users issuing symptom-related queries to a search engine (could be own symptoms, or asking on behalf of someone else)</td>
<td>General (Queries related to any symptoms from a comprehensive list of symptoms from the Merck medical dictionary)</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White &amp; Horvitz (2013a)</td>
<td>Unclear, participants were users issuing symptom-related queries to a search engine (could be own symptoms, or asking on behalf of someone else)</td>
<td>General (Queries related to any symptoms from a comprehensive list of symptoms from the Merck medical dictionary)</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yastik (2017)</td>
<td>Symptomatic; patients presenting to ED with acute symptoms</td>
<td>General (any acute symptoms, no acute distress related to airway, breathing, or circulation)</td>
<td>Not assessed</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.3.4 RQ1: What proportion of different populations use the Web to appraise symptoms?

Four studies, all surveys, reported the proportion of the study sample that engaged in Web use for symptom appraisal (Table 2.3).

In Fox & Duggan’s (2013) population-based survey with adults living in the USA, 35% reported going online to attempt self-diagnosis. Participants were sampled to mirror the population in terms of demographics, but disproportionately-stratified to increase the incidence of non-White respondents. This survey was conducted in the USA and is therefore likely to reflect proportions in Western, high-income countries with high Internet penetration.

White & Horvitz’s (2009b) survey conducted among Microsoft employees found that “three quarters of subjects” (the authors do not provide absolute numbers; assuming the proportion is 75%, the 95% CI for N = 515 is 71% to 79%) reported searching for information on symptoms. “Two thirds” reported researching professionallyundiagnosed conditions, at least once a month (White & Horvitz, 2009b). It should be noted that this sample was biased towards younger, male respondents with high educational level and socio-economic status working within an industry that is very Web-oriented.

Medlock et al. (2015) examined online health information seeking in older people by surveying members of a senior Christian organisation. They found that 23% of participants reported using the Web in the past 12 months to determine the cause of symptoms. This shows that, while Web use for symptom appraisal may be less common among older people than in the general population (compared to 33 – 37% found by Fox et al. (2013)), older people do engage in it.

While the previous surveys focused on diagnostic searches for any conditions/symptoms, Thomson et al. (2012) conducted a survey with colorectal cancer patients and found that 25% of the sample reported pre-diagnosis Web use for symptom appraisal.
Table 2.3. Percentage of people engaging in Web use for symptom appraisal reported by included studies.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study population</th>
<th>Sample size</th>
<th>Percentage of the sample who report Web use for symptom appraisal</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fox &amp; Duggan (2013)</td>
<td>Adults living in the USA</td>
<td>3,014</td>
<td>35%</td>
<td>33-37%</td>
</tr>
<tr>
<td>White &amp; Horvitz (2009b)</td>
<td>Microsoft employees, USA</td>
<td>515</td>
<td>75%</td>
<td>71-79%</td>
</tr>
<tr>
<td>Medlock et al. (2015)</td>
<td>Members of a senior church organisation, Netherlands</td>
<td>118</td>
<td>23%</td>
<td>15-31%</td>
</tr>
<tr>
<td>Thomson et al. (2012)</td>
<td>Colorectal cancer patients, USA</td>
<td>242</td>
<td>25%</td>
<td>20-31%</td>
</tr>
</tbody>
</table>

**Conclusion RQ1**: Fox et al.’s (2013) study with its large, population-based sample size is most likely to give an accurate proportion for the general population (in a western, higher income country), though the other included studies give an indication of how this proportion can vary depending on the population being surveyed, i.e. depending on socio-demographic variables and disease-related factors. It should be noted, however, that all studies were conducted several years ago, and given the rapidly increasing penetration of the Internet (Office for National Statistics, n.d.), findings may not accurately reflect current proportions.

**2.3.5 RQ2: Which symptoms are likely to be researched online?**

Seven studies examined characteristics of symptoms that were searched for online (De Choudhury et al., 2014; Fiksdal et al., 2014; Hay et al., 2008; North et al., 2011; Rice, 2006; Simons et al., 2017; Thomson, Siminoff, & Longo, 2012). Three of these were survey studies (Hay et al., 2008; Rice, 2006; Thomson, Siminoff, & Longo, 2012), two were interview studies (Fiksdal et al., 2014; Simons et al., 2017) and two involved analyses of data on usage of online resources like social media, search engines and health websites (De Choudhury et al., 2014; North et al., 2011).

North et al. (2011) compared users of the Mayo Clinic website with people who used a telephone triage system to appraise their symptoms, and found that telephone triage users were more likely to have acute and conspicuous symptoms requiring immediate relief, while Website
users were more likely to research chronic conditions. Hay et al. (2008) surveyed 120 rheumatology patients and found that some individuals in their study sought help online because they had a history of undiagnosed symptoms. North et al.’s and Hay et al.’s findings both suggest symptoms are researched online when they have been present for a prolonged time. However, a study with patients presenting to the Emergency Department shows that individuals may also research symptoms when they have been present for less than twelve hours (Yastik, 2017).

In their study with 242 colorectal cancer patients, Thomson et al. (2012) found that neither symptom severity nor stage at diagnosis was related to Web use for symptom appraisal, but Web users were more likely to experience symptoms typically perceived as embarrassing, such as a change in bowel habits. Similarly, Choudhury, Ringel Morris & White (2014), who analysed log data obtained from Twitter and a search engine found that while symptoms were generally tweeted more than searched, potentially embarrassing, stigmatised or sensitive symptoms such as ‘vaginal bleeding’ or ‘pelvic pain’ were more likely to be searched for than tweeted.

Furthermore, Rice’s (2006) population-based telephone survey conducted in the US concluded that more frequent online health seekers were more likely to look for sensitive health topics which are difficult to talk about than less frequent online health seekers.

Finally, in Fiksdal et al.’s (2014) focus group study with 19 US adults, participants reported turning to the Web when symptoms were perceived as trivial/non-serious and they wanted to avoid ‘bothering’ health professionals.

**Conclusion RQ2:** It appears Web use for symptom appraisal occurs when symptoms are persistent, have a history of being undiagnosed by health professionals, are potentially embarrassing or stigmatised, and/or when they are perceived as superficial/non-serious.

### 2.3.6 RQ3: How is Web use for symptom appraisal conducted (search strategies)?

**Theme 1: Symptom-based, condition-based and treatment-based searches**

Three distinct approaches to searching were identified: (i) symptom-based searches, which used symptoms as search terms; (ii) condition-based searches, which involved searches for particular conditions, and (iii) treatment-based searches which involved researching treatments for symptoms without prior research on possible causes.
Log data from search engines suggest the majority (65%) of exploratory health-related searches, i.e. those aimed at diagnosing a condition, are symptom-based rather than condition-based (Cartright et al., 2011), and remain symptom-based throughout the search, as search sessions tend to start and end with purely symptom-related queries (Cartright et al., 2011; White & Horvitz, 2009a). One should bear in mind, however, that log-based studies cannot ascertain searchers’ actual intentions and motivations. The authors assume occurrences of certain search terms signal certain intentions (e.g. a symptom and the term “cause” signals diagnostic intent), however, papers did not report any prior validation of these algorithms.

An experimental study which observed 79 people as they used Google or a symptom-checker tool to diagnose symptom vignettes reported that most users conduct symptom-based searches, as most people began their search by entering symptoms, and only 24% began by specifying a condition (Luger et al., 2014).

In an experimental study conducted by Perez et al. (2015), 78 participants were instructed to research the Web as if they were experiencing a given symptom, described in a vignette, while using Think Aloud. Think Aloud (also known as ‘cognitive interviewing’) requires participants to vocalise their thoughts while performing a task (Lewis, 1982). The authors found that 19% of searches were treatment-based, and the remainder symptom or condition-based (the authors did not report these separately). Participants in another vignette-based study also reported treatment-based searches (Simons et al., 2017). It should be noted that the external validity of vignette-based studies is limited, as individuals base their searches on the vignette descriptions rather than actual perceptions or observations, and the search behaviour observed is likely to depend on the phrasing of the vignette.

Keselman, Browne and Kaufman (2008) used interview and Think Aloud methods to explore how a convenience sample of 20 lay individuals interpreted a symptom vignette using the American consumer health information service MedlinePlus. They concluded that some participants conducted condition-based searches, and some participants used a symptom-based approach. Additionally, they identified a group of participants who used a condition-based approach, but began their search with a broader hypothesis, such as “heart disease” and then attempted to narrow down their search. Different barriers seemed to play a role in the different search strategies: Those conducting condition-based searches were prone to confirmation bias,
seeking out information that confirmed their hypothesis, and terminating the search before reviewing further hypotheses. Those starting with a broad hypothesis often terminated the search without coming to a conclusion. The symptom-based searchers struggled to find relevant results due to the lack of specificity of their search terms (Keselman et al., 2008).

Overall, it seems most searches are symptom-based, as both log-based studies, which have high external validity, and experimental studies, which have high internal validity, confirm this finding. The algorithms used for the log-based studies were not validated, however, and experimental and qualitative studies used to observe search behaviour have limited generalisability to real-world contexts.

**Theme 2: Selection of search terms**

Keselman et al. (2008) examined the search behaviour of 20 adults using Think Aloud, and discovered that participants often ignored symptoms mentioned in the vignette if they perceived them as irrelevant and exempted these from their search terms. The authors termed this ‘selective perception bias’. The participants also tended to ignore aspects of duration of the symptoms and had difficulty discerning acute from chronic symptoms. However, these findings stem from a single study with only 20 adults using fictional scenarios, and thus further validation is required.

**Theme 3: Age differences**

Two studies reported on age differences in search behaviour (Chin & Fu, 2010; Chin et al., 2009). Chin (2009) compared the search performance of 69 younger and older adults while performing either a well-defined task (searching for a specific medical term on a website) or an ill-defined task (using the website to diagnose a set of symptoms). The study found that older adults performed better in the ill-defined task, while younger participants performed better in the well-defined task (Chin et al., 2009).

In another study, Chin (2010) presented 46 older and younger adults with different interfaces of the same website; one interface categorised symptoms based on the body parts they occurred in, whereas the other interface categorised symptoms according to functional systems (e.g. respiratory system). Younger adults tended to click on significantly more links within one category, suggesting they followed the interface of the website, while older adults clicked significantly more between-category links regardless of the interface. Chin and Fu (2010)
conclude this was due to older adults using their existing medical knowledge rather than the interface to guide their search, which is supported by their finding that older adults performed better in a medical knowledge task (Chin & Fu, 2010).

Thus, there are some indications that older adults perform differently in Web searches for symptom appraisal than younger adults, possibly due to medical knowledge. The available studies, however, used small sample sizes, thus inferences to the wider population may not be appropriate.

Theme 4: Selecting information

Several studies examined how users select information from their search results. I identified four sub-themes relating to selection of information.

Number of search results viewed

Lauckner and Hsieh (2013) report that participants in their laboratory-based, experimental study with 274 undergraduate students viewed approximately four links on results pages, which was the number of results visible above the fold, i.e. users would need to scroll down to view more results. Corroborating this finding, Keselman et al. (2008) discovered that participants in their qualitative study (N=20) often ignored relevant links while trying to diagnose a symptom vignette if these were located below the fold. Luger et al. (2014) found that older adults (N=79) in their experimental, lab-based study tended to view approximately 6 conditions on the WebMD symptom checker tool after entering a set of symptoms, though the authors do not clarify whether this was the number visible above the fold.

Thus, the top results returned by search engines will have maximum impact on symptom appraisal, whereas those located below the fold may have little to no effect. As these findings all relate to lab-based studies, however, further investigation in naturalistic settings would be beneficial.

Process of elimination

In their study using Think Aloud with 79 adults aged 50+ years, Luger et al. (2014) found that 91% of participants used a ‘process of elimination’, whereby the symptoms described in the vignette were compared against those listed for a given condition, and the condition was
discarded as a hypothesis if it included symptoms not mentioned in the vignette. This finding suggests that a common search strategy is to narrow down the hypotheses by discarding those with symptoms not matching one’s own. However, the majority of the sample was highly educated (all had some University education), therefore generalisability to the wider population is unclear.

**Source credibility**

In Luger et al.’s (2014) study with 79 adults aged 50+, source credibility was mentioned by only 25% of the sample. They also found that a third viewed user-generated content like discussion boards, which are not quality-controlled. White & Horvitz (2013a) found in their log-based study that consumer sites like MayoClinic.com or WebMD (both well-known American corporations and health websites) are positively related to click-through rates following searches which contain symptoms as search terms, suggesting widely known, established health websites are likely to be accessed during Web use for symptom appraisal. White & Horvitz (2009a) also suggest searches which ‘escalate’ (i.e. progress from searching for symptoms to serious conditions) contain more visits to ‘trusted sources’ (e.g. governmental websites, websites of health organisations). It is important to note that, while the authors claim to research ‘diagnostic searches’, this was identified through the presence of symptoms in search terms and may therefore also include non-diagnostic searches.

**Pages mentioning serious illnesses**

White and Horvitz (2013a) filtered logs from a search engine, and examined how users issuing symptom queries subsequently interacted with search results pages. They found that users engage more with captions on search results pages that mention serious illnesses, hovering more frequently and longer over these captions and clicking these more often than captions mentioning benign causes. Terminology related to serious illnesses such as “malignant”, “severe”, and “tumour” significantly increased click probability, while terms like “benign” decreased click probability. Additionally, users are more likely to engage with sites indicating they can help identify causes of symptoms, e.g. by mentioning the words “learn” and “causes”.
While we do not know searchers’ intentions or how they used the information found, these findings suggest those researching symptoms online are more likely to engage with websites relating to serious causes.

**Conclusion RQ3:** Web use for symptom appraisal typically involves inputting information into a search tool, and subsequently narrowing down results returned by the search tool. When inputting information, most users appear to choose search terms based on symptoms rather than hypothesised conditions, but users do not appear to utilise all information available (e.g. some symptoms may be omitted, as well as the frequency/duration of symptoms). Furthermore, there is some limited evidence that older adults perform differently in Web searches for symptom appraisal than younger adults, and that this may be due to medical knowledge. Once a selection of results is provided by the search tool, users tend to narrow results down by taking into account the results’ position on the results page, the degree of seriousness of the condition, the credibility of the source, and the extent of overlap between the listed and the experienced symptoms.

**2.3.7 RQ4: Behavioural consequences of Web use for symptom appraisal**

**Theme 1: Increased help-seeking**

In Fox and Duggan’s (2013) population-based survey (N=3,014), 46% (95% CI 44.22% to 47.78%) of online self-diagnosers claimed that information found online led them to think they needed the attention of a health professional. Thomson et al. (2012) found in their survey with colorectal cancer patients that 25% (95% CI 19.54% to 30.46%) of online self-diagnosers were reportedly persuaded by the information found online to see a health professional. This suggests the proportion of people encouraged to seek medical help based on Web use for symptom appraisal may be significantly lower among colorectal cancer patients than in the general population (Fox & Duggan, 2013).

Using logistic regression with a survey sample of over 2000 Americans aged 12+, Ybarra et al. (2006) found that online self-diagnosers were 2.5 times more likely to report contacting a health professional than online health information seekers who did not try to diagnose symptoms online, suggesting Web use for symptom appraisal is linked to increased healthcare contact.

One study suggests that the mode of presenting information on a website may affect users’ decisions to seek medical advice: In a UK-based qualitative study (Cumming et al., 2010),
participants reportedly felt encouraged to seek help after viewing an online ‘digital storytelling’ video about urogenital atrophy.

Using log-based search engine data, White and Horvitz (2010b) examined search behaviours related to healthcare utilisation intent. Healthcare utilisation intent was assumed to be present when users conducted searches for healthcare practitioners/clinics near their geographical area. They found that users who displayed certain search behaviour - such as visiting websites that mention serious conditions before benign ones – were more likely to show healthcare utilisation intent subsequently. This suggests online search behaviour following symptom queries is related to subsequent healthcare contact, though mechanisms of causality are unclear based on this data. Furthermore, the authors do not report whether/how their algorithms were validated; it is unclear whether their proxy measure of healthcare utilisation intent is valid.

By observing how patients attending a rheumatology clinic completed the NHS and WebMD symptom-checker tools, Powley et al. (2016) found indications that symptom checker tools provide information that can propagate unnecessary help-seeking. Of 34 patients, 15 were inappropriately advised to seek help from emergency services rather than primary care, indicating potential issues with the algorithms used in symptom-checker tools. One should bear in mind, however, that this study does not allow any conclusion on whether real users would follow this advice or not.

**Theme 2: Decreased help-seeking**

In Powell et al.’s (2011) interview study with 26 users of the NHS Direct website (the official website of the UK National Health Service), some participants reportedly used online health information as a form of ‘demand management’, to identify trivial symptoms not warranting medical attention. Similarly, some participants in Fiksdal et al.’s (2014) focus group study reported using the Web to avoid ‘bothering’ health professionals with trivial symptoms. Finally, in Attfield et al.’s (2006) interview study, participants reported sometimes being reassured by Web searches that help-seeking was not necessary. None of the studies followed up whether users had correctly or incorrectly classified symptoms as trivial. Furthermore, evidence for this theme stems only from qualitative studies, thus generalisability is uncertain.
Theme 3: Communication with health professionals

Fox et al. (2013) found that 53% of online self-diagnosers reportedly discussed the health information found online with a health professional. Three qualitative studies (Attfield et al., 2006; Cooper et al., 2013; Simons et al., 2017) found that patients used the Web to appraise symptoms in order to prepare for consultations, by preparing questions, collating relevant information and enhancing their knowledge in order to understand the advice received.

Two studies found indications that Web use for symptom appraisal is related to reduced communication with a health professional (Hay et al., 2008; Thomson, Siminoff, & Longo, 2012). Hay et al. (2008) found that new rheumatology patients who engaged in Web use for symptom appraisal were significantly less likely to want to challenge their health professional’s advice than those who did not. The authors note that study participants were concerned about evoking the impression of questioning health professionals’ advice. Thomson et al. (2012) found that Web use for symptom appraisal was significantly related to feeling hesitant about discussing symptoms with a health professional. Direction of causality is unclear, however. It is possible that information found online dissuaded individuals from communicating with health professionals, or that people chose to research symptoms online because they were reticent about discussing their symptoms.

Conclusion RQ4: From the above findings, we can conclude that Web use for symptom appraisal is used to inform the decision of whether to present to health services, and that online self-diagnosers are more likely than other health information seekers to contact a health professional. This can potentially be increased, where appropriate, with novel methods like ‘digital storytelling’. Some evidence also suggests that online health information can potentially reduce help-seeking by calming users’ fears. It is unclear, however, what proportion of users feel encouraged or discouraged to seek help appropriately (i.e. what proportion of users who feel encouraged to seek help actually have a condition warranting medical attention, and what proportion of users who feel discouraged to seek help actually do not need medical attention).

Furthermore, it is unclear whether those engaging in Web use for symptom appraisal are more or less likely to seek medical advice than those experiencing the same symptoms without researching online, as this comparison was not made in any of the included studies. Web use for symptom appraisal can also play a role in communication with health-professionals by
influencing how individuals prepare for consultations, and prompting discussion of online health information.

2.3.8 RQ5: Emotional consequences of Web use for symptom appraisal

In White & Horvitz’s (2009b) survey among 515 Microsoft employees, 38.5% reported that online health information had made them feel anxious in the past, and 50.3% reported Web use for symptom appraisal had made them feel less anxious. As might be expected from the sampling frame, the survey sample was biased towards younger, more educated and IT-literate respondents.

Powell et al. (2011), who examined the motivations of users of the NHS Direct website using semi-structured interviews, found that participants sought health information online to obtain reassurance about symptoms. The majority nevertheless subsequently sought medical help, though sometimes with less urgency and anxiety.

Teriaky, Tangri & Chande (2015) surveyed patients (N=87) awaiting appointments at a general gastroenterology clinic and asked those who reported using the Web prior to their consultations whether this had changed their anxiety levels. 77% experienced no change, 21% experienced an increase and 2% a decrease. One should note that this sample consists of those who decided to report to health services; there may be a larger proportion of users who felt calmed by their searches and therefore did not present to health services.

Lauckner et al. (2013) found in their experimental laboratory-based study (N=274) that presenting search engine results relating to serious conditions before benign conditions, and a higher frequency of results relating to serious conditions, was related to negative emotional outcomes such as fear. These findings suggest a causal relationship between exposure to search results during Web use for symptom appraisal and increases in anxiety.

Another experimental study conducted by Norr et al. (2014), however, found no difference in the anxiety levels of their participants (N=56) following review of either a) websites containing information on causes of symptoms or b) websites on general health and wellness (e.g. exercise, healthy diet).

Conclusion RQ 5: Some evidence suggests there is a relationship between Web use for symptom appraisal and health anxiety. Findings from experimental studies were mixed
regarding causal relationships. Surveys and interviews indicate there is a potential for decreases in anxiety and calming effects, and that the proportion who report feeling calmed by Web use for symptom appraisal is higher than those reporting anxiety. It is also possible that those who engage in Web use for symptom appraisal are more anxious about their health generally. It is unclear when anxiety is warranted as participants’ actual diagnoses were not followed up, and comparisons to those who did not research symptoms was lacking.

2.4 Discussion

This is the first systematic review and synthesis of the literature available on Web use for symptom appraisal. My main findings are:

- About 35% of the general population engage in Web use for symptom appraisal, but the proportion can vary considerably (25-75%) depending on the population under study (given the rapidly increasing penetration of the Internet (Office for National Statistics, n.d.), however, findings may not accurately reflect current proportions)
- Symptoms tend to be researched online when they are long-term, potentially embarrassing/stigmatised, have been presented to health services previously with inconclusive outcomes, and/or when they are perceived as trivial
- Searches tend to be based on symptoms rather than hypothesised conditions; users seem to focus on particular symptoms, while disregarding other symptoms and aspects like frequency and duration
- Once a selection of results is returned by the search tool, people use specific techniques to narrow results down, e.g. taking into account the position on the results page, or the credibility of the source
- Evidence indicates that online information is used to inform the decision of whether to contact health services, and is related to (increased and decreased) anxiety, but the precise impact cannot be discerned due to lack of follow-up and appropriate comparison groups.

One of the aims of this PhD project is to explore whether the Web is a facilitator or barrier to help-seeking, and to explore options of leveraging the Web to promote earlier presentation to health services. Below I therefore discuss whether Web use for symptom appraisal should be
viewed as an asset or a liability in healthcare delivery based on currently available evidence, and make recommendations for the improvement of online health information.

2.4.1 Web use for symptom appraisal – assistance or hindrance to health promotion?

Criticisms of online self-diagnoses include concern over unnecessary anxiety and healthcare utilisation (White & Horvitz, 2009a, 2013b). This review confirms that the Web can increase anxiety and healthcare contact among users (Lauckner & Hsieh, 2013; White & Horvitz, 2009a, 2009b, 2013b), but reveals that there is insufficient evidence to conclude this occurs unnecessarily.

First, it is important to note limitations of approaches used to examine relationships between Web use for symptom appraisal and health anxiety or help-seeking behaviour. Cross-sectional surveys cannot show causality. It is possible that using the Web to appraise symptoms causes anxiety, or that anxiety triggers Web use for symptom appraisal, or that a third factor influences both. Furthermore, the surveys which reported on anxiety among online self-diagnosers were biased towards certain demographic (White & Horvitz, 2009b) or patient groups (Teriaky et al., 2015), and did not use validated measures of anxiety levels.

Log-based studies, which evaluate behaviour based on search engine log data, do not allow firm conclusions regarding users’ actual behaviours and motivations. For example, White & Horvitz (White & Horvitz, 2009a) found that users who begin their searches for symptoms often progress to researching serious conditions, but it is not clear whether users are anxious or using the information to reassure themselves. The authors assume certain search terms signal certain intentions (e.g. a symptom and the term ‘cause’ signals diagnostic intent), but no action was described to determine the validity of these assumptions. Thus insights from log-based studies are limited.

Experimental research shows that users asked to research certain symptoms may report feeling anxious following Web searches (Lauckner & Hsieh, 2013), but it is not possible to infer whether this anxiety would be unwarranted in a naturalistic setting, i.e. if symptoms were actually present.

Using the Web to appraise symptoms may also decrease anxiety in some cases (Powell et al., 2011; White & Horvitz, 2009b) and Web searches are sometimes used to identify alternatives to healthcare utilisation (Attfield et al., 2006; Powell et al., 2011). Individuals describe using online
information to evaluate mild/superficial symptoms, in order to avoid wasting health professionals’ time (Fiksdal et al., 2014). This shows that Web use for symptom appraisal can also decrease anxiety and help-seeking. It is possible that Web use for symptom appraisal discourages help-seeking for trivial symptoms, thus reducing pressure on healthcare resources. However, it is also possible that Web use for symptom appraisal leads to complacency and prevents help-seeking when it is actually necessary.

There are also indications that Web-based information can help individuals recognise their symptoms as signs of serious conditions (Thomson, Siminoff, & Longo, 2012). Indeed, this review highlights that online health information is an important resource when there are (perceived) barriers to obtaining information from health professionals, e.g. when symptoms are embarrassing or stigmatised (De Choudhury et al., 2014; Rice, 2006; Thomson, Siminoff, & Longo, 2012), or when previous visits to healthcare have been ineffective (Hay et al., 2008; North et al., 2011). This suggests there is potential for the Web to be an assistance to healthcare.

Finally, it should be noted that worry can also have positive effects on health behaviours (Chae, 2015). The Protection Motivation Theory suggests that fear will increase intention to perform a certain behaviour if the individual feels able to perform the behaviour, and believes that the behaviour will reduce the threat (Conner & Norman, 2015). Similarly, according to the Health Belief Model, fear should result in recommended health behaviour, if perceived benefits of the recommended behaviour are high and barriers are low. Overall, this indicates anxiety induced through online health information can enhance recommended health behaviours, if information is presented in a way that enables concrete action and decision-making.

A limitation I discovered across different methodologies was the lack of follow-up on participants’ help-seeking behaviour and diagnoses. Without this information, we cannot discern whether individuals’ self-diagnoses and decisions regarding help-seeking behaviour are appropriate, or not. We also cannot determine long-term impacts on healthcare utilisation. Furthermore, essential comparison groups are generally lacking. For example, it would be necessary to compare those who research symptoms online with those who do not (rather than surveying only online self-diagnosers), and to compare those who present to health services
with those who do not (rather than surveying only patients presenting in clinic), in order to
determine impacts of Web use.

2.4.2 Recommendations to improve online health information

Based on the findings of this review, I suggest changes to health websites, Web applications
and search engines such that they can provide useful information to those researching
symptoms.

My analyses reveal that users tend to search inductively based on symptoms. Search engines
and symptom checker tools need to ensure users are directed to useful information when
symptom terms are entered. The review also shows that searchers tend to omit dimensions like
duration and frequency of symptoms in their search terms (Keselman et al., 2008), and that
symptom-based searchers struggle to find relevant results due to the lack of specificity of their
search terms. This suggests it is important that users are directed to useful terms to narrow their
search and prompted to provide information on duration and frequency of symptoms, in order to
improve specificity of searches. In support of this, recent research suggests incorporation of
query expansion techniques into information retrieval systems can improve the effectiveness of
search engines for diagnostic symptom searches (Palotti et al., 2015).

This review also reveals that online health information can impact on the decision to seek help,
and on communication with health professionals. Health websites and applications need to
ensure they provide useful information to support searchers in their decisions and healthcare
interactions. Health websites providing symptom information could, for example, provide clear
guidelines on when medical advice should be sought (e.g. if a symptom has a certain quality or
duration), how help should be sought (e.g. immediately via emergency services or within the
next week via primary care), and possibly where help can be sought based on location
information.

2.4.3 Strengths and Limitations

As the review includes a diversity of study types and methods, a quantitative synthesis or meta-
analysis was not possible. However, traditional forms of systematic review which do not make
use of all forms of evidence often do not take differing contexts into account, limiting their use to
policy-makers and practitioners (Dixon-Woods et al., 2005). Combining findings from different
study designs allows a richer, more holistic understanding of the phenomenon under study
I was able to combine real-world insights from observational studies, such as analyses of search engine log data, with data from more controlled, experimental settings, thereby improving external and internal validity. Furthermore, by incorporating findings from large, population-based studies as well as smaller interview-based studies, I was able to gain an understanding of the impacts of Web use for symptom appraisal at population level while also obtaining more detailed reports of peoples’ perceptions and experiences. Moreover, by including studies covering a broad range of populations (different conditions/symptom profiles, age groups, socioeconomic status etc.) I have shown how Web-usage can differ depending on context.

In this review I considered a diversity of symptoms and conditions; when more research in this area becomes available, it would be useful to carry out more focused reviews, as the nature of the symptom is likely to influence Web use online (De Choudhury et al., 2014; Fiksdal et al., 2014; Hay et al., 2008; North et al., 2011; Rice, 2006; Thomson, Siminoff, & Longo, 2012).

Finally, it should also be noted that this review did not examine Web use for mental health symptoms. Web use for symptoms related to mental health, and its impacts on help-seeking, also represents an important field of study and should be assessed in a separate review of the literature.

### 2.4.4 Conclusions and future work

This systematic review indicates that the Web can disseminate information to those worried about symptoms and can impact on their decisions to present to health services. It also suggests Web use for symptom appraisal can impact on how patients prepare for consultations with healthcare professionals. Thus, we can conclude that Web use for symptom appraisal has the potential to influence the timing of help-seeking, and the communication between patients and healthcare professionals during consultations.

At present, limitations of the reviewed studies mean it is not clear when the Web plays a beneficial role in healthcare delivery, and when it is detrimental. Web use for symptom appraisal has been linked to increased, as well as decreased anxiety and healthcare contact. However the evidence does not show when this is warranted, because studies did not follow up whether participants ultimately sought help following their Web searches, and whether they received a
diagnosis. Furthermore, comparison groups are lacking to determine impacts of Web use for symptom appraisal.

2.5 Web use for symptom appraisal and lung cancer

So far, I have discussed Web use for symptom appraisal in relation to a wide variety of conditions. I will now discuss how the findings presented here relate to the specific case of lung cancer and this PhD project.

2.5.1 Are people with lung cancer likely to engage in Web use prior to diagnosis?

Research suggests Web use for symptom appraisal is related to sociodemographic factors (Fox & Duggan, 2013; Thomson, Siminoff, & Longo, 2012). One study found an association with female gender, younger age, White ethnicity, higher income, and higher education level (Fox & Duggan, 2013). As lung cancer is associated with male gender, older age, Ethnic minority group, and low SES (Torre, Siegel, & Jemal, 2016), it is plausible to assume Web use for symptom appraisal among people with lung cancer will be low. However, one should bear in mind that family and friends access online health information on behalf of patients pre-diagnosis (Thomson, Siminoff, & Longo, 2012), and family members play an important role in help-seeking for lung cancer symptoms (Braybrook, Witty, & Robertson, 2011; Corner et al., 2006; L. K. Smith et al., 2005; Tod et al., 2008). Furthermore, Web use is increasing in the 65+ population (Zickuhr & Madden, 2012), suggesting the importance of this information source will increase in future older generations.

As discussed above, individuals are more likely to research symptoms related to chronic conditions online than acute illnesses (North et al., 2011), which suggests the Web may be an important information source for people with (lung) cancer symptoms. Moreover, symptoms perceived as embarrassing or stigmatised are also likely to be researched online. Due to the link between smoking and lung cancer, people with lung cancer tend to experience stigmatisation and blame (Chapple, Ziebland, & McPherson, 2004) and sometimes avoid seeking medical advice due to fear of stigmatisation (Tod et al., 2008). Thus individuals experiencing symptoms they perceive to be related to their smoking behaviour may turn to the Web for preliminary information and advice.
Additionally, my synthesis showed that individuals with a history of undiagnosed symptoms tend to engage in Web use for symptom appraisal (Hay et al., 2008; Morgan et al., 2014). Patients with lung cancer are more likely than other cancer patients to present to primary care multiple times before suspected cancer referral (Georgios Lyratzopoulos, Neal, Barbiere, Rubin, & Abel, 2012).

Thus, Web use among people with lung cancer may be low due to the demographic profile of the patient group. However, an increasing importance can be expected among future generations as Internet use is increasing, especially as the characteristics of the illness (chronic, stigmatised symptoms which often remain undiagnosed for a prolonged time) suggest it is likely to be researched online prior to diagnosis.

### 2.5.2 Are people with lung cancer likely to find useful / relevant information?

The review suggests that online diagnozers tend to research symptoms rather than conditions, and tend to focus on one particular symptom, often ignoring dimensional aspects like duration and frequency (Cartright et al., 2011; Keselman et al., 2008; Luger et al., 2014; White & Horvitz, 2009a). This means that people with lung cancer may struggle to find relevant information, as the symptoms are unspecific and can be related to a wide range of other conditions, particularly if dimensional aspects are not taken into account. For example, simply researching 'cough' will yield mostly results relating to common cold whereas researching ‘persistent cough’ or ‘cough for more than 3 weeks’ will yield more specific results. Additionally, users tend to view only the search results appearing at the top of the screen, and do not tend to scroll downwards to view subsequent results (Chin & Fu, 2010; Keselman et al., 2008; Lauckner & Hsieh, 2013; Luger et al., 2014). If unspecific search terms are used, it is less likely that relevant information will appear at the top of search results.

Finally, users seem to use a process of elimination to determine whether a given condition is a possible cause for symptoms, comparing their symptoms against those listed and discarding conditions for which the match is low. This search behaviour may be problematic in lung cancer, as symptoms vary in type and number and individuals often experience only one or two symptoms. For example, the NHS lung cancer symptoms website lists 15 symptoms for lung cancer (NHS Choices, n.d.-g). An individual experiencing only one symptom (e.g. breathlessness) is therefore likely to (falsely) discard this hypothesis.
2.6 Implications for this PhD project: What are the gaps and how will I address them?

From the above, I conclude that online health information is likely to play an increasingly important role in pre-diagnosis information seeking among people with lung cancer in future. However, whether users are likely to access useful information and whether the Web will play a positive, facilitating role or act as a barrier to help-seeking, is unclear from the literature. Research suggests how information is presented is important, but ideal forms of presentation are unknown. In the following I therefore aim to establish:

i) What role does Web-based information currently play in help-seeking prior to a diagnosis of lung cancer?

ii) How can health information available online be improved to encourage early presentation to health services?

The first aim will be addressed using a mixed-methods exploratory approach (Chapter 3), while the second aim will be addressed using an experimental factorial design to determine causal impacts of different forms of information presentation on help-seeking intentions (Chapters 4-7).
Chapter 3: The role of Web-based health information in help-seeking behaviour prior to a diagnosis of lung cancer: A mixed-methods study

The material presented in this chapter has been published in (Mueller, Jay, Harper, & Todd, 2017) but is partly rewritten and expanded upon for this thesis.

3.1 Introduction

As the systematic review of the literature presented in the previous chapter highlights, the Web has the potential to play a role in the help-seeking process. Evidence indicates Web-based information can impact on decisions to present to health services, and how individuals communicate with their healthcare providers. The findings of the review suggest that the proportion of people with lung cancer researching their symptoms online is likely to be low. The review also showed that there is insufficient evidence to conclude whether the Web is most likely to act as a facilitator or barrier to help-seeking.

In order to meet this project’s aim of assessing the role of the Web prior to a lung cancer diagnosis and leveraging this to encourage earlier presentation, I first needed to establish whether people with lung cancer actually access this information source, and how they perceive its effects. This chapter seeks to explore these questions.

3.1.1 The pathway to diagnosis

As described in the introductory chapter of this thesis, the route from symptom recognition to diagnosis and commencement of treatment has been theorised into four cyclical intervals by Walter et al. (2012) in a model of pathways to treatment (Figure 1.2). The length of these intervals can be influenced by disease factors (e.g. site, size growth rate), healthcare provider and system factors (e.g. access to resources, healthcare policy), and patient factors (e.g. psychosocial factors).

Here I focus on the three intervals leading up to diagnosis (the appraisal, help-seeking, and diagnostic interval, Figure 1.2), as low survival rates in lung cancer have been linked to delays to diagnosis (Holmberg et al., 2010; Radzikowska et al., 2012). I refer to these three intervals as the ‘pathway to diagnosis’.
I focus on patient factors, as research has shown that people with lung cancer often experience symptoms for several months before presenting to health services (Corner et al., 2005, 2006; Tod et al., 2008). Research suggests lack of knowledge about lung cancer symptoms is one of the biggest barriers to help-seeking (Corner et al., 2005, 2006; McCutchan et al., 2015; McLachlan et al., 2015; Shim et al., 2014; L. K. Smith et al., 2005; Tod & Joanne, 2010). Furthermore, symptoms are often masked by pre-existing comorbidities that have similar symptoms, making it difficult for the patient to distinguish between existing and new symptoms (Corner & Brindle, 2010; Tod & Joanne, 2010). Fear of being diagnosed with cancer and fatalistic beliefs about treatability of lung cancer may also impede help-seeking (Tod et al., 2008; Tod & Joanne, 2010).

Lung and colorectal cancer patients have been shown to be proactive when appraising symptoms (McLachlan et al., 2015) and health-related Web use has been documented in various cancer populations such as lung, colorectal, prostate, testicular, breast, cervical, and bowel cancer (Peterson & Fretz, 2003; Thomson, Siminoff, & Longo, 2012; Ziebland et al., 2004). This suggests the Web could play a role in the time before cancer diagnosis, e.g. if people with cancer search the Web for information to appraise their symptoms. While evidence indicates that people with lung cancer do access the Web (Peterson & Fretz, 2003; White & Horvitz, 2016), the proportion of Web users is likely to be low, because lung cancer patients tend to be older (> 70 years) and to have lower education levels and socio-economic status (Torre et al., 2016), and these factors are related to low levels of health-related Web use (Fox & Duggan, 2013). Overall, there is a growth of health information on the Web and an increasing tendency for individuals to seek health information online (Chou et al., 2011). However, little is known about how people make use of this source prior to diagnosis, as most research focuses on Web usage after patients have been diagnosed.

Due to the scarcity of previous research on this topic, an exploratory approach was required to gain a preliminary understanding of the potential role of online health information during the time leading up to a lung cancer diagnosis. In this study, I aimed to gain this preliminary understanding by exploring patients’ own retrospective accounts of how they remember the events leading up to their diagnosis, with particular focus on the perceived impact of Web searches on this process. Previous research has shown that family members sometimes conduct Web searches on behalf of patients (Thomson, Siminoff, & Longo, 2012) and that
family members play an important role in lung cancer patients’ help-seeking behaviour (Braybrook et al., 2011; Corner et al., 2006; L. K. Smith et al., 2005; Tod et al., 2008). Therefore, I also aimed to explore accounts of proxies (family/friends of patients), and whether they assisted patients with online searches or conducted searches on their behalf.

3.1.2 Aims and research questions

My overall aim was to gain a preliminary, exploratory insight into whether Web-based information plays a role in the pathway to lung cancer diagnosis. To meet this aim, I addressed three research questions:

Q1: Which lay information sources do people with lung cancer (or proxies) retrospectively report using to appraise symptoms prior to diagnosis, and what proportion retrospectively report researching their condition online?

Q2: In cases where pre-diagnosis Web searches take place, how do individuals perceive the impact of the information they find on their pathway to diagnosis?

Q3: What are possible barriers to using the Web pre-diagnosis?

3.2 Method

3.2.1 Design

My research questions required the combination of quantitative and qualitative methods: Quantitative methods were used to establish the proportion of lung cancer cases in which pre-diagnosis Web searches took place (Q1). Qualitative methods were used to explore individuals’ perceptions of the impact their Web searches had on the pathway to diagnosis, as well as barriers that might prevent individuals from accessing the Web for health information pre-diagnosis (Q2+3). Finally, mixed methods were required because a survey was needed to screen for relevant individuals for interview, as I expected low levels of Web use among lung cancer patients.

Thus, this study comprised 1) a cross-sectional, retrospective survey of recently diagnosed lung cancer patients attending hospital clinics and 2) a qualitative interview study with a purposive sub-sample of the survey participants.
3.2.2 Participants and recruitment

I recruited recently diagnosed lung cancer patients to explore patients’ retrospective accounts of the events leading up to their diagnosis. Participants were recruited from outpatient clinics at two large university hospitals in the Northwest of England between July 2014 and March 2015. Patients were eligible if they (i) had received a lung cancer diagnosis in the six months prior to study entry, (ii) had sufficient English language to complete the questionnaire, (iii) were able to consent, and (iv) reported experiencing at least one symptom before diagnosis. Patients whose diagnosis was more than six months before study entry were excluded to reduce recall bias (Thomson, Siminoff, & Longo, 2012).

Participants were sampled for qualitative interviews purposively based on questionnaire responses regarding (i) whether the Web had been used prior to diagnosis and (ii) sociodemography (age, gender, smoking status) as these sociodemographic factors have been shown to be related to Web use (Valero-Aguilera et al., 2012). I included both Web users and non-Web users to gain insight into reasons for and against using the Web prior to diagnosis. Proxies were invited to participate in interviews, as they tend to be involved in health information-seeking (Schook et al., 2014), and have been shown to play an important role in lung cancer patients’ help-seeking (Braybrook et al., 2011; Corner et al., 2006; L. K. Smith et al., 2005; Tod et al., 2008). I recruited proxies who engaged in Web searches prior to diagnosis on the patient’s behalf, and those who did not. Data collection continued until saturation was reached, i.e. when no new evidence for theoretical points emerged and I began to note similar accounts recurring (Spencer et al., 2013).

3.2.3 Procedure

Ethical approval was granted by the NHS Health Research Authority NRES Ethics Committee (Ref:14/SW/0128, IRAS Project ID 143412, approval letter provided in Appendix 5). I approached patients attending outpatient clinics, who had previously been identified from clinical notes as potentially eligible, and provided them with the participant information sheet (Appendix 6). Following informed consent, I provided participants with a questionnaire and stamped-addressed return envelope, with the option of completing the questionnaire in clinic (with the researchers, if they wished) or at home. A subset of consenting participants was selected purposively to participate in follow-up interviews, which were conducted in clinic.
### 3.2.4 Measures

**Questionnaire**

The paper-based questionnaire (Appendix 7) took 10-15 minutes to complete. Questions were standardised and assessed:

(i) Which symptoms were experienced before the diagnosis

(ii) Length of time between first noticing symptoms and presenting to health services (on an ordinal scale: 1 week, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, more than 6 months)

(iii) Perceived reasons for delay (participants were asked “What thoughts did you have about these symptoms before you sought medical help?” and were asked to select from a list created from themes identified in previous studies on help-seeking for lung cancer (Birt et al., 2014; Brindle et al., 2013; Corner et al., 2005, 2006; Tod et al., 2008; Tod & Joanne, 2010; Walton et al., 2013), e.g. “I was afraid that it might be something serious.”

(iv) Whether the patient and/or a proxy had used any lay information sources prior to diagnosis to help understand the symptoms/condition (Internet/Web, leaflets, magazines, books, television, radio, medical journals or libraries, public library, other),

(v) Rating of information sources used (very helpful, helpful, unhelpful, very unhelpful)

(vi) Details on Web searches conducted prior to diagnosis if applicable (who conducted the search, search engine and search terms used, websites accessed),

(vii) Information on habitual Web/technology use (whether the Internet is ever used, if yes, number of hours during a typical week) and

(viii) Sociodemographic information (age, sex, education level, and employment status).

The development of the questionnaire was informed by previous literature on help-seeking behaviour and Web searching behaviour (Peterson & Fretz, 2003; Qureshi, Morkve, & Mustafa, 2008; Valero-Aguilera et al., 2012), medical reference works (Bourke & Burns, 2011; Dilworth & Baldwin, 2001), brainstorming within the research group, as well as discussion with a Patient and Public Involvement group for cancer and palliative care.
Interviews

Interviews were semi-structured with open-ended questions and standardised prompts (Appendix 8). The interview topic guide covered:

- Symptom experience prior to diagnosis, with a focus on motivators and barriers to seeking help
- Web searches conducted prior to diagnosis (if applicable), with a focus on perceived impacts on the pathway to diagnosis (e.g. the decision of whether to present to health services)
- Reasons for and against using the Web prior to diagnosis.
- Use of other lay information sources, including talking to family/friends and TV/radio (mass-media lung cancer symptom awareness campaigns)

Clinical records

Following consent, type of lung cancer and smoking status were obtained from patient records.

3.2.5 Analyses

Quantitative analysis

I analysed questionnaire data descriptively using IBM SPSS V.22 to calculate percentages, means and standard deviations. For proportions, I calculated 95% confidence intervals as an indication of the variability of the results and to facilitate comparisons using the Wilson method (Brown et al., 2001).

Qualitative analysis

Interviews were audio-recorded, transcribed verbatim and organised using QSR NVivo10. Framework Analysis (Spencer et al., 2013) was used to identify recurring and important themes in the data. My analysis involved the following five stages (Tierney et al., 2011; Ward, Furber, Tierney, & Swallow, 2013):

1. familiarisation with the data

Familiarisation was achieved by repeatedly reading all interview transcripts, and noting recurring topics.
(2) development of a theoretical framework

A broad framework of topics was developed to organise the data, based on the interview protocol as well as recurring topics identified in step (1). The topics were then sorted and grouped under broad categories to create a hierarchical structure of topics and sub-topics.

(3) indexing data

The framework was then applied to the data by using NVivo to label transcript sections according to the topics occurring in each section. This was undertaken by at least two independent researchers, and any discrepancies discussed until consensus was reached.

(4) summarising data in thematic charts

A matrix was created within NVivo10 (QSR International Pty Ltd. Version 10, 2012) for each topic, with participants in the rows and sub-topics in the columns. Transcript sections were then summarised into the relevant cells, keeping as close to participants’ original wording as possible. To illustrate, an excerpt from a framework matrix is provided in Appendix 9.

(5) synthesising data by mapping and interpreting

Matrices were next explored by comparing cells across participants and within participants to identify similarities or differences in how participants described their experiences. This facilitated identification of recurring themes and links between themes. Themes were discussed in the research group until consensus was reached. To aid interpretation, I categorised participants’ reported Web searches according to the processes in the pathway to diagnosis (Walter et al., 2012) they reportedly impacted on, according to interviewees’ accounts and perceptions, i.e. appraisal of symptoms (appraisal interval), decision to consult (help-seeking interval), and healthcare professional appraisal (diagnostic interval).

Epistemology and ontology. While qualitative research paradigms such as ethnography and grounded theory tend to be rooted in specific epistemological and philosophical stances, Framework Analysis should be viewed as a data analysis method rather than a research paradigm and it is thus more flexible in terms of epistemological and ontological underpinnings (Ward et al., 2013). In terms of ontology, Framework Analysis is most closely associated with subtle realism, which assumes that an objective reality exists outside of human beliefs and understanding, but that it can only be known through socially constructed meanings (Ormston,
Spencer, Barnard, & Snape, 2014; Ward et al., 2013). This stance is clearly distinguished from naïve realism, which assumes that an objective reality exists and can be perceived and depicted accurately (Ormston et al., 2014). It also stands in contrast to idealism, which asserts that an objective reality does not exist outside of the human mind and socially constructed meanings (Ormston et al., 2014). From a subtle realist perspective, qualitative research aims to understand the external reality, but remains aware that this can only be interpreted through the perceptions of research participants and researchers. In terms of epistemology, Framework Analysis assumes a mostly inductive position with a bottom-up approach, building knowledge and developing theory through observations (Ormston et al., 2014). Unlike approaches like Grounded Theory which are entirely inductive and allow no influence of prior knowledge, theory or findings from previous literature on the development of new theory, Framework Analysis does account for the influence of existing knowledge on the analytic process (Ward et al., 2013).

3.3 Survey results

3.3.1 Sample description

Between July 2014 and March 2015, 199 patients attending the two hospital outpatient clinics were identified as eligible, and 122 consented (61.3%). Nine participants were excluded after consent as they had not experienced symptoms prior to diagnosis (it was not possible to discern this in advance from clinical records). Thus 113 participants were included in the final sample (Figure 3.1). The mean age was 67 (SD=8.8) years, and ranged from 42 to 88 years. The majority were male (56.6%), retired (67.3%), former smokers (73.6%), and reported educational attainment below University level (93.9%); 55% had non-small cell lung cancer (Table 3.1). Approximately half of all participants reported having an Internet connection at home (51.3%) and 61.1% reportedly had used an Internet device at some point. 23% did not own any Internet device (Table 3.2).
199 patients identified as eligible

122 patients consented and completed the survey

113 patients included in final survey

88 patients consented to be contacted for an interview

17 patients and 16 proxies purposively sampled for interviews (24 interviews conducted)

N = 77 did not consent

Reasons:
Not interested: n = 8
Felt too tired: n = 6
Did not want to fill in study documents: n = 4
Felt too anxious / overwhelmed: n = 9
Was handed a study information pack but did not return the questionnaire: n = 23
Did not attend their appointment: n = 11
Patient was not approached by clinician: n = 3
Reason missing: n = 13

9 excluded as survey showed they had not experienced symptoms prior to diagnosis (this was not apparent from clinical records)

Figure 3.1 Participant recruitment flow diagram (STROBE diagram).
<table>
<thead>
<tr>
<th><strong>Table 3.1 Participant demographic details.</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Employment status</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Education level</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Type of lung cancer</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Smoking status</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Table 3.2 Habitual Web/technology use among participants.

<table>
<thead>
<tr>
<th></th>
<th>n (% out of N=113)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet connection at home</td>
<td>58 (51.3%)</td>
</tr>
<tr>
<td>Use of a computer</td>
<td>60 (53.1%)</td>
</tr>
<tr>
<td>Owning a computer</td>
<td>73 (64.6%)</td>
</tr>
<tr>
<td>Use of a Web-enabled mobile phone</td>
<td>24 (21.2%)</td>
</tr>
<tr>
<td>Owning a Web-enabled mobile phone</td>
<td>40 (35.4%)</td>
</tr>
<tr>
<td>Use of a tablet</td>
<td>34 (30.1%)</td>
</tr>
<tr>
<td>Owning a tablet</td>
<td>38 (33.6%)</td>
</tr>
<tr>
<td>Ever use any Internet devices</td>
<td>69 (61.1%)</td>
</tr>
<tr>
<td>Not owning any Internet devices</td>
<td>26 (23.0%)</td>
</tr>
</tbody>
</table>

When asked which thoughts they had had about their symptoms before the diagnosis, most participants endorsed ‘I thought my symptoms were related to another condition’ (72/112, 64.3%). More than a third reportedly did not think they were ill at first (40/112, 35.7%), and 31.3% (35/112) reportedly thought they could self-treat at home. One in five (23/112, 20.5%) stated that they did not think their symptoms required medical attention. Some were reportedly afraid that their symptoms indicated a serious illness (20/112, 17.9%), and 13.3% (15/112) acknowledged that they had not wanted to bother their doctor. Some patients reported difficulty in obtaining an appointment with their GP (12/112, 10.7%). Only 8% (9/112) attested that they had felt uncomfortable seeing their GP due to an unhealthy lifestyle. The self-reported time interval from first symptom perception to contacting a healthcare provider ranged from 1 week to ‘more than 6 months’ (Figure 3.2), with a median interval of one month between symptom perception and contacting health services.
3.3.2 Use of lay information sources

Approximately half (60/112, 53.6%, 95% CI=44.5%-62.5%) reported that either they or a proxy had used a lay information source to appraise their symptoms prior to diagnosis. Almost half reportedly used some form of lay information source themselves (53/112, 46.9%, 95% CI=38.3%-56.5%), and 26.5% (30/110, 95% CI=19.8%-36.3%) disclosed that a proxy had used a lay information source on their behalf. Figure 3.3 depicts which sources participants reported using either themselves or by proxy.

The most commonly used lay information source for appraising symptoms was ‘talking to family or friends’, which was endorsed by 38.7% (43/111, 95% CI=29.7%-47.8%) of participants. Talking to family or friends about symptoms was rated as ‘somewhat’ or ‘very helpful’ by 26 of the participants who used this source (26/33, 10 missing). Twelve (12/112, 10.7%, 95% CI=6.2%-17.8%) participants reported that they, or a proxy, used information from television to evaluate their symptoms, and three reportedly used the radio (3/112, 2.7%, 95% CI=0.9%-7.6%). Information from TV was rated as ‘somewhat’ or ‘very helpful’ by 9 of participants, for the other 3 this data was missing. Information from the radio was rated as ‘somewhat helpful’ by all
three participants. None of the participants reported use of leaflets, magazines, books, medical journals, or public libraries to appraise their symptoms. None of the participants named any other lay information sources apart from those listed (although the survey provided a space for this).

Figure 3.3 Use of lay information sources to appraise symptoms prior to diagnosis by participants and family/friends ("proxies").

3.3.3 Proportion of people with lung cancer who report Web searches prior to diagnosis

20.5% of the sample (23/112, 95% CI=13.0%-28.0%) reported they, or a proxy, researched symptoms online prior to diagnosis. Seven (7/112, 6.3%, 95% CI = 1.8% to 10.8%) stated they researched their symptoms themselves, four of these with the help of a proxy. Family/friends conducted online searches on the patient’s behalf in a further 16 cases (14.3%, 95% CI= 7.8%-20.8%), thus the majority of searches (20/23) involved proxies. While 23 people reported the Web was used in their case, some reported several Web users, e.g. a spouse and grandchild. Thus, in total, I identified 31 Web users: 7 patients; 7 spouses/partners; 12 sons/daughters; 2 sons-in-law/daughters-in-law; 1 grandchild; 1 nephew; and 1 friend. Fifteen participants (15/20) rated the information they encountered online as ‘very helpful’ or ‘somewhat helpful’, while 5/20 rated it as ‘very unhelpful’ or ‘somewhat unhelpful’ (3 missing).
3.3.4 Description of Web searches

Of the 23 participants who reported Web searches prior to diagnosis, 20 reported that Google was used to search; the rest did not know which search engine was used. The majority (19/23) reported using the NHS Direct website. Other websites included WebMD (7/23), patient.co.uk (5/23), Yahoo Health (2/23), and Netdoctor (1/23). Two participants reported visiting discussion forums.

Twenty-one participants reported search terms used (Table 3.3). Eight participants used symptoms as search terms (e.g. “persistent cough”), and five used possible causes/conditions such as “throat cancer” or “stopping smoking”. Three people used investigative test results that had been communicated to them by health professionals before they had received a final diagnosis (e.g. “pleural effusion”). In these three cases, the Web searches took place before a diagnosis was given, but after some investigation of symptoms had been initiated. Five participants used combinations of these (e.g. “stomach cancer and weight loss”). Overall, five participants reported that “lung cancer” was included in their search.

To summarise, only a small proportion of the surveyed patients reported use of the Web prior to their diagnosis to appraise their symptoms/condition. In most cases, searches were conducted by or with the help of a proxy. In semi-structured interviews, I explored patients’ and proxies’ views and experiences in more detail; this is presented in the following section.
Table 3.3 Search terms used during Web searches prior to diagnosis.

<table>
<thead>
<tr>
<th>Category</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms</td>
<td>Persistent cough (x4)</td>
</tr>
<tr>
<td></td>
<td>Hoarse voice, hoarseness, croaky voice (x3)</td>
</tr>
<tr>
<td></td>
<td>Back pain, shoulder pain, lower back pain (x2)</td>
</tr>
<tr>
<td></td>
<td>Coughing up blood (x1)</td>
</tr>
<tr>
<td></td>
<td>Lump on neck (x1)</td>
</tr>
<tr>
<td></td>
<td>Breathlessness (x1)</td>
</tr>
<tr>
<td></td>
<td>Swollen face and neck (x1)</td>
</tr>
<tr>
<td></td>
<td>Recurrent chest infection (x1)</td>
</tr>
<tr>
<td></td>
<td>Pain in chest (x1)</td>
</tr>
<tr>
<td></td>
<td>Weight loss (x1)</td>
</tr>
<tr>
<td></td>
<td>Night sweats (x1)</td>
</tr>
<tr>
<td></td>
<td>Dry mouth (x1)</td>
</tr>
<tr>
<td>Conditions / possible causes</td>
<td>Lung cancer, lung cancer symptoms (x5)</td>
</tr>
<tr>
<td></td>
<td>Cancer (x1)</td>
</tr>
<tr>
<td></td>
<td>Throat cancer (x1)</td>
</tr>
<tr>
<td></td>
<td>Stomach cancer (x1)</td>
</tr>
<tr>
<td></td>
<td>Myasthenia (x1)</td>
</tr>
<tr>
<td></td>
<td>Stopping smoking (x1)</td>
</tr>
<tr>
<td></td>
<td>Anxiety (x1)</td>
</tr>
<tr>
<td>Test results (communicated to the patient before the final diagnosis was given)</td>
<td>Pleural effusion, fluid in lungs (x1)</td>
</tr>
<tr>
<td></td>
<td>Patch on vocal cords (x1)</td>
</tr>
<tr>
<td></td>
<td>Iron anaemia (x1)</td>
</tr>
<tr>
<td></td>
<td>Collapsed lung (x1)</td>
</tr>
</tbody>
</table>

3.4 Qualitative interview results

Twenty-four qualitative interviews were conducted; ten with patients, seven with proxies and seven with the patient and proxies together. In total, 33 people were interviewed (n=19 female). Fourteen interviews involved Web searches prior to diagnosis, either by the patient or proxies.
3.4.1 Perceptions of the impact of pre-diagnosis Web searches on the pathway to diagnosis

In the following, I group Web searches according to the intervals in Walter et al.’s (2012) model. Within each interval, I explore participants’ perceptions of how their Web searches impacted on the processes described in the model.

**Appraisal interval: Perceived impacts on appraisal and self-management**

Some participants reportedly accessed the Web for information once they had perceived bodily changes, but had not yet decided to present these to a healthcare professional. Online information was used to identify possible causes of symptoms.

“And [husband’s name] being how he is, he won’t go to the doctors anyway, so we did sort of self-diagnose, if you like.” [R27, wife of patient, 50-64 years]

Some participants described that the information they read online about symptoms changed the way they appraised their symptoms, causing them to view symptoms as more serious than before, and in some cases even convincing them that the cause was lung cancer.

“I just put in, to start with, shoulder pain, and lung cancer came up straight away. […] And that’s…I thought, oh, you know; I looked at it, and I thought, lung cancer? Crikey! Because I’d no idea that people got pain anywhere near there. And so I went on one website after another, after another, after another, just to try and read the symptoms to see if the symptoms were all the same on each site, and they were, basically.” [R18, daughter of patient, 51-65 years]

One participant described that the information she found online caused her to view her husband’s symptoms as less serious. In this particular case, the searcher entered the symptoms experienced by her husband as search terms, coupled with her hypothesis of what was causing the symptoms, i.e. smoking cessation. For example, she reportedly searched for “stopping smoking and cough”. Using this search strategy, she reportedly felt reassured by the information she found that the symptoms were caused by smoking cessation rather than a disease.

R27: “You see, the sweating all night and the coughing. We’d had a look online and his friends had told him, he’d stopped smoking. So that happens, you get insomnia, you can’t sleep either and you’re coughing a lot. And you’re just bringing anything up that’s been in your lungs for years. So we sort of left that at that, thinking that’s what it was. […]”

**Interviewer:** “And in your search, did you, at any point, come across any information about lung cancer?”
R27: “I don’t think I did […] my specific search terms were, stopping smoking, so I kept putting stuff with smoking in. I didn’t put night sweats, if you like, and get the whole…the amount of what would cause it, if you like.”

[…] Interviewer: “Could you just tell me a little bit more about how you went about your search?” […]

R27: “So I probably put, stopping smoking and night sweats. And then what’s come up about that. Stopping smoking and cough…how long after stopping smoking, will they carry on coughing?” [R27, wife of patient, 50-64 years]

Help-seeking interval: Perceived impacts on the decision to consult a healthcare professional

According to Walter et al.’s (2012) model, individuals in the help-seeking interval form the decision to consult a healthcare professional and make an appointment, and the interval is concluded when a first consultation takes place. Several participants in my interview study reportedly used online information to inform their decision on whether to present to a healthcare professional.

“I kept thinking, this cough’s not clearing. But like I said, it went on months and months […] so that’s when my son went on the internet, and that’s when he said, ‘mum, Aunt [name] and Aunt [name] they’re right, you need to go.’ And that’s when I went.” [R22, patient, >65 years]

Some participants who conducted Web searches before first consultation with a healthcare professional perceived no impact on their decision-making, because they had reportedly already formed the decision to present to health services before they began their search.

“I knew there was something wrong, that you had to go and see a doctor […] The decision was made before I even googled it, yes.” [R4, patient, 51-65 years]

Diagnostic interval: Perceived impacts on healthcare professionals’ appraisals

In Walter et al.’s (2012) model, the diagnostic interval commences following first consultation and involves appraisal by a healthcare professional, investigations, referrals and appointments.

Some participants described presenting to health services multiple times, without diagnostic tests to determine the cause of the symptoms. Participants reportedly turned to the Web because they felt dissatisfied with the advice they received from healthcare professionals.

“So I went on the internet, I think because she’d had four visits to the doctor and we weren’t getting anywhere, so I went on to just see, you know, if I could find anything out really to give me an idea what else it could be other than an allergic reaction.” [R14, daughter of patient, 51-65 years]
Interviewees reportedly used the information found online to challenge their doctors’ advice by suggesting other possible causes for symptoms and requesting further tests. Participants felt that their assertiveness in urging further investigation impacted on healthcare professionals’ decisions to conduct diagnostic procedures.

“So then I went online and I put in facial swelling and neck swelling and it said it could be an infection of the glands or the ducts. So that’s why I asked the GP, when I went back with her on the fifth visit, could it be an infection in the glands or the ducts. [...] I’m not saying he wouldn’t have done it but I think the fact that I was with my mum and maybe being a little bit more assertive instigated him to maybe look a little bit further. Yeah, definitely. [...] I wasn’t rude but I was assertive, and it was only then that he investigated further and listened more closely to her chest.” [R14, daughter of patient, 51-65 years]

3.4.2 Perceived barriers to using the Web pre-diagnosis

I also interviewed patients and proxies who had not accessed the Web prior to diagnosis, in order to understand perceived barriers to pre-diagnosis Web use. This can help to understand whether the role of the Web in the pathway to lung cancer diagnosis may change in future.

 Concern over unnecessary worry and fear

Of those who reported not researching their condition online prior to diagnosis, several reportedly avoided this because they were concerned that it could lead to unnecessary worry and fear.

“Sometimes it can frighten the life out of you, you know what I mean? It’s like when people used to buy the home medical directory and you’d got a headache and something else, when you looked it up, you’ve got everything under the sun.” [R8, husband of patient, >65 years]

 Preferring not to know

Some participants stated that they wanted to know as little as possible about their health, preferring to leave decisions to health professionals.

“No I don’t like looking it up. I don’t really like knowing unless I’ve got to.” [R21, patient, 51-65 years]

 Believing symptoms trivial

One participant felt her symptoms were too mild and familiar to warrant further research.

“I mean I like to know what’s going on but I wouldn’t research a tickly cough because I’ve been in that situation many times before so…” [R6 patient, female patient, >65 years]
Unfamiliar with or not interested in technology

Most of those who did not research their condition online were not comfortable using technology.

*Interviewer: “Do you ever use the computer?”*

*R24: “No, because I can’t even, it takes me all this time to text and reply. I prefer a conversation, you know, ring somebody. You know, further than that, I just make a mess of everything.” [R24, patient, 65+ years]*

Some expressed disinterest in the use of technology.

“*I am computer illiterate… and I prefer it that way. Yeah. I used to use a computer when I worked, it’s not that I can’t, it’s that I’m not interested; it’s such a waste of time.”* [R6, patient, 51-65]

A few participants wanted to access the Internet, but lacked the skill.

“The lads have got it now and I think they, I think they are brilliant. I wish it’s one of the things, I wish I could, but I’ve never gone onto the Internet.” [R21, patient, 51-65 years]

3.4.3 Search strategies and information appraisal

Interviewees described how they went about their online research. Based on the Framework Analysis conducted, I identified three recurring themes.

**Process of elimination**

Interviewees described selecting information using a “process of elimination”, by first establishing a list of possible causes for the symptoms, and then eliminating or discarding hypotheses that did not match their symptom profile.

“Yeah because I looked at lower back pain and there was a lot of information it could have been lots of different things. Erm, you know, some more serious than others so I really went through a process of elimination. I think there was one article where it showed a lot of red flag indicators some of which were evident, so she didn’t have a cough that brought up blood so I really kind of for my research thought actually it’s probably, I don’t know what it is but it might not be lung cancer because there was some you know red flag indicators that it wasn’t.” R11 [Daughter-in-law of patient, 35-50]

**Trust in known national organisations**

Interviewees reported particular trust in the NHS website and other known UK organisations such as the charity Macmillan Cancer Support. Several interviewees expressed distrust towards American websites, as significant differences between health care systems were perceived.

“I think I’d always probably go to the National Health website because that’s where you get the best information because you can get conflicting information if you go on an
American site. It might be just as pertinent but you don’t know that they’re using the same treatment and the same things.” R26 [Patient, >65 years]

**Checking for congruence**

Interviewees described visiting several websites and comparing information across websites. If several websites provided the same information, this was seen as a sign for the trustworthiness, reliability and credibility of the information.

“And so I went on one website after another, after another, after another, just to try and read the symptoms to see if the symptoms were all the same on each site, and they were, basically.” R18 [Daughter of patient, 35-50 years]

“…so there’s literally, all sorts of things, but once you read it a few times on different websites, you’re like, oh well, I think it’s just this and this.” R27 [Wife of patient, 51-65 years]

**3.4.4 Use of other lay information sources**

Although this PhD project focuses on use of the Web, it is nevertheless important to explore other lay information sources as well, to help assess what information individuals find useful and why. As the survey results show, participants reported using several other lay information sources apart from the Web to help them appraise their condition prior to diagnosis. Most reported talking to proxies about their symptoms, and some also reported using information from TV/radio. Below, I use interview data to elucidate how participants perceived and utilised these information sources.

**The Web and proxies’ role in help-seeking**

Involvement of proxies emerged as one of the most important facilitators to seeking medical help. Many relatives took on the role of urging the patient to seek help or return to the doctor.

“But it was only my family kept pushing me, really. They kept saying, ‘you need to go. You need to go’. After me son said, ‘Mum, please go’ he said ‘there’s something not right’. I did [inaudible] and he sent me for an X-ray and that’s when it all started.” R22 [Patient, >65 years]

In some cases relatives took on a more active role. Rather than simply urge the patient to seek help, they accompanied them to consultations, asked questions of the doctors and requested further tests.

“Dad got quite stressed because he knew something was wrong and his wife then rang us up on Saturday night and went please help, something’s not right. Dad’s not right. The doctors have said there’s nothing wrong with him and they gave him some antidepressants but she knew there was something wrong. Dad knew there was something wrong. […] So I went to ask him [GP] for a gastroscopy and he agreed on what we did. I took a photo of my dad in March when he was on his cruise with his
trunks and he had full weight and healthy and when he saw the picture he was quite shocked and he understood how thin he was at that point.” R2 [Daughter of patient, 35-50 years]

In some cases, proxies identified symptoms the patients themselves had not noticed.

“[Name of wife] noticed a change in a cough that I had, a persistent cough, which I probably wouldn’t have responded to, I hadn’t really noticed the change.” R4 [Patient, 51-65 years]

**Media campaigns**

Some participants commented on their experiences with media campaigns on lung cancer awareness, which ran in the UK in the time preceding the present study. Fifteen interviewees recalled hearing ads on television or radio prior to being diagnosed, but only three interviewees reported some impact on decision-making (two of these were proxies).

“And I don’t get the chance to listen to the radio but in my GP’s surgery there’s big posters and maybe subconsciously that actually prompted me to keep on at her to get it sorted.” [R28, daughter of patient, 35-50 years]

One interviewee, a patient’s spouse, reportedly used the information from the campaign to support her claims to the patient that help-seeking was warranted.

“So I was like, please, you’ve got to go to the doctors now, because you are coughing all night. And I think that, along with the TV thing about, if you stop smoking and you’re still coughing, then it’s time to see a doctor. I think that the two of them at the same time, must have forced him to go.” [R27, wife of patient, 51-65 years]

The third interviewee reported that she found the ads useful as they gave her a specific behavioural target.

“I think it was very informative and I think it gives you the target, doesn’t it, to do something. A lot of the time a cough for three weeks is not a lot because you could have a really bad cold and cough, cough, cough for three weeks and then vanish but I think it’s an indicator, good indicator to do something about it certainly if it’s recurring.” [R26, female patient, >65 years]

However, several interviewees reported that the ads were less useful for the time following the first visit to the GP.

“I mean that was one of the things, wasn’t it, I mean I kept saying, well, I’ve done that, I’ve been back to the doctor, you know. [...] Not much you can, you can only go back so many times till the doctors say come on, you know.” [R26, female patient, >65 years]

Twelve interviewees reported no impact of the media campaigns on their decision-making. Apart from the relevance of smoking and coughing, interviewees seemed to be unaware of other symptoms and risk factors.

I: “And do you remember anything else, or seeing anything on TV about lung cancer?”
"Only that smoking causes lung cancer. But I don’t smoke and never have, so, you know, I don’t take a lot of notice of that – not really." [R18, daughter of patient, 35-50 years] 

Furthermore, many interviewees reportedly ignored the campaigns because they felt they were not relevant to them. People seemed to be optimistically biased into assuming the ads would not apply to them.

"Yes I think it was and I just put it down to a bit of a cough, you know, just a chest cough or something, I just put it down to that. You don’t think, you know, even though you have got a cough and there is advertisements on the telly and you don’t think, well I don’t take it all in. I’m thinking, well that’s not me, you know, I’m fine, I’ll be fine – and then I wasn’t." [R21, female patient, 51-65 years]

3.5 Discussion

This is the first study to explore pre-diagnosis Web searches among lung cancer patients. I found that approximately a fifth of the sample of lung cancer patients reported pre-diagnosis Web searches to research symptoms and help them understand their condition, with most searches conducted by proxies. Furthermore, my analyses showed that patients and proxies perceived impacts of their pre-diagnosis Web searches on their pathways to diagnosis, including symptom appraisal, forming the decision to seek help, and interactions with healthcare professionals.

My overall aim was to gain a preliminary, exploratory insight into whether Web-based information plays a role in the pathway to lung cancer diagnosis. To explore this role, I discuss below i) the proportion of people with lung cancer reporting pre-diagnosis Web searches ii) perceived impacts of the Web searches on the pathway to diagnosis and iii) what prevents people from accessing the Web, and whether this is likely to change in future.

3.5.1 Proportion of people with lung cancer reporting pre-diagnosis Web searches

I found that 20.5% of my sample reported Web searches prior to diagnosis to help appraise symptoms or understand their condition. The majority of searches were conducted by or with the help of a proxy. Although more than half (61.1%) of the patients in the survey indicated having used the Internet in the past, and approximately half (51.3%) reported having an Internet connection at home, only 6.3% (95% CI = 1.8% to 10.8%) of patients reported researching their condition online themselves.
Only one other published study has examined the proportion of cancer patients who engage in pre-diagnosis Web searches: In a study with colorectal cancer patients, Thomson et al. (2012) found that 25% (95% CI = 19% - 31%) of patients had researched symptoms online themselves, not including patients whose family or friends searched on their behalf. Comparing my 1.8-10.8% with Thomson et al.’s 19-31%, the proportion found in my sample of lung cancer patients is clearly lower. This may be due to my participants being older with lower education levels than those in Thomson et al.’s (2012) study, as these factor have been related to lower levels of health-related Web use (Fox & Duggan, 2013). Differences in symptoms profiles between colorectal and lung cancer may also play a role (Mueller, Jay, Harper, Davies, et al., 2017).

3.5.2 Perceived impacts of Web searches on the pathway to diagnosis

In my qualitative interview study, I explored how patients and proxies perceived the impact of pre-diagnosis Web searches on the events leading up to diagnosis. By mapping participants’ accounts onto the model developed by Walter et al. (2012), I showed that participants perceived an influence of the information they found online on processes in all three intervals leading up to diagnosis (appraisal, help-seeking, diagnostic).

**Symptom appraisal (appraisal interval)**

Participants reportedly used online information to assess the seriousness of their symptoms and to identify possible causes. Previous research has suggested a rising importance of online health information on symptom appraisal processes; in the US, for example, more than a third of adults report having used online information to identify the cause of symptoms (Fox & Duggan, 2013).

My findings suggest that participants searching in the appraisal interval had differing experiences with Web-searches, with some reporting that the information led them to believe their symptoms were serious, whereas others reported that the information reassured them that symptoms were not serious.

My interview findings tentatively suggest that search strategies may play an important role in how online information affects appraisal of symptoms. The participant who reportedly felt reassured had conducted a hypothesis-driven search strategy, by researching symptoms paired together with her hypothesised cause, ‘stopping smoking’. Search results were therefore biased towards the hypothesised cause. Previous research (Keselman et al., 2008) has suggested that
searchers who use hypothesis-driven searches are prone to certain forms of bias, such as confirmation bias (starting with a hypothesis and confirming it), and premature termination bias (stopping after viewing only one topic). Further research on the differential effects of symptom-driven and hypothesis-driven searches is necessary to determine how search functions on health websites should be designed to enhance patients’ ability to appropriately appraise symptoms.

**Decision to consult (help-seeking interval)**

Several participants reported that online information was used to inform their decision of whether to present to health services, and several participants reported that it encouraged them to make an appointment with a healthcare professional. Previous research confirms that most ‘online diagnosters’ subsequently seek a professional medical opinion (Fox & Duggan, 2013; Ybarra & Suman, 2006). In Thomson et al.’s study with colorectal cancer patients, about a quarter of patients reportedly felt persuaded by the information found online to see a health professional. Furthermore, analyses of search engine log data have indicated that those who research symptoms online often show healthcare utilisation intent subsequently, e.g. by searching for clinics near their geographical area (White & Horvitz, 2013b).

Overall, previous research coupled with my results suggest that there may be a causal relationship between Web use and deciding to seeking help; this should be examined quantitatively in future research. Research in this area would be especially crucial for conditions like lung cancer, where earlier presentation to health services can maximise chances of survival.

**Healthcare professional appraisal, investigations, referrals and appointments (diagnostic interval)**

My analysis suggested that the Web can play a role in the diagnostic interval. This finding was particularly interesting because it was unexpected. I had expected the Web to play a role in the time from first symptom perception until presentation to health services, because this interval has been termed the ‘patient interval’ suggesting this is the interval during which patient characteristics and behaviour are relevant (Weller et al., 2012). Also, Safer et al.’s model of patient delay suggests information-seeking would occur in this interval (Safer et al., 1979). Furthermore, previous campaigns have focused on this interval and on encouraging
presentation to health services rather than interactions with healthcare professionals following first contact (Ironmonger et al., 2014). This finding from my interview study highlights that information-seeking and patient factors remain relevant during the diagnostic interval.

My findings suggest Web-based health information can empower patients and their families to appraise and challenge doctors’ advice and request further diagnostic procedures. This is particularly interesting because efforts to reduce patient delays to diagnosis in lung cancer have focused on encouraging presentation to health services (Cancer Research UK, n.d.-a; Ironmonger et al., 2014; Tustin, 2012), i.e. the appraisal and help-seeking interval. Little attention has been paid to the role patients and proxies play in the diagnostic interval.

Recent years have seen a shift towards patient-driven health care, with patients increasingly interested in decision-making (Pelzang, 2010). Web-based health information has been associated with this shift, by increasing patients’ awareness of health professionals’ fallibility and uncertainties in diagnoses (A. G. H. Thompson, 2007). Thus, with the help of Web-based health information, patients may play an increasingly important role during the diagnostic interval. This is particularly important because my results indicate that individuals turn to the Web when they are dissatisfied with advice received from healthcare professionals and when they experience delays in obtaining a diagnosis. If future research and interventions focus on how the Web can be utilised to support patients in this role, delays to diagnosis may be reduced.

3.5.3 Barriers to pre-diagnosis Web searches: Current and future importance of the Web prior to a lung cancer diagnosis

I explored barriers to using the Web for health information prior to diagnosis. This can help to assess the extent of the role the Web can play in lung cancer patients’ pathways to diagnosis, and whether this role is likely to change in future.

The majority (79.5%) of my sample reported not accessing the Web prior to diagnosis, for a range of reasons. While some of these barriers, such as perceived triviality of symptoms, may persist in the future, others are likely to change. For example, while some participants in my study reportedly preferred to defer to the doctor and not know details regarding their own health, engagement in healthcare is increasing and patients generally desire more detailed information from health professionals than they receive (Al Qadire, 2014; van Weert, Bolle, van Dulmen, &
Jansen, 2013). Furthermore, participants reported unfamiliarity with, or disinterest in technology. This is unsurprising as the current cohort of those aged 65 and over, who are at highest risk for lung cancer, are less likely than any other age group to access the Internet (Zickuhr & Madden, 2012). Future lung cancer patients will be more familiar with the Web (Chevalier, Dommes, & Marquié, 2015; Couper et al., 2010).

3.5.4 Recommendations and future research

The association between pre-diagnosis Web searches and length of intervals in the pathway to diagnosis should be assessed statistically with a larger sample size using a validated, reliable measure of time intervals leading up to diagnosis. To date, no fully validated measure exists (for example Neal et al.’s measure (2014) is only validated for the reported time of first GP consultation). Measures used in previous research have considerable limitations (R. Andersen et al., 2009), or were too lengthy for use in the present study (e.g. Neal et al., 2014), as I wished to keep the screening survey brief to reduce the burden for patients and to enable a sufficiently large sample to be screened for interview. The present study, as an exploratory study in this field, focused on a qualitative exploration of perceptions and reported experiences, and was not designed to deliver a precise quantitative assessment of the association between Web use and time to diagnosis.

Furthermore, strategies to utilise the Web to encourage early presentation to health services should be investigated in future research. For example, my analyses indicate that individuals turn to the Web when they experience difficulty communicating with health professionals. Future research should therefore explore how health websites can provide information that will help facilitate patients’ communications with health professionals.

My results suggest hypothesis-driven searches (search terms based on hypothesised conditions) and evidence-driven searches (search terms based on symptoms only) may have differing effects on how individuals evaluate symptoms. Future research should systematically investigate differential effects of search strategies on symptom appraisal and subsequent help-seeking behaviour. This will help to inform the development of Web-based symptom appraisal tools and search engine algorithms.
3.5.5 Limitations

My findings relied on retrospective, self-reported measures of patients, pertaining to events that took place prior to diagnosis. Cancer patients’ reports of the events leading up to diagnosis can be inconsistent (Hess et al., 2012). Patients in my study had been diagnosed up to 6 months prior to study entry, and were asked to recall events before the diagnosis. Therefore some of my measures may be subject to recall bias. For example, patients may not have been able to recall all search terms used during their pre-diagnosis Web searches. It was not feasible, however, to identify and recruit individuals with lung cancer prior to diagnosis.

Furthermore, patients may have been unaware of Web searches conducted by proxies so this variable may be under-reported in the survey, as the survey was completed by patients. This could be addressed in future research by conducting a large-scale survey among patients and proxies.

Finally, as with all qualitative research, one must be cautious in generalising from my purposive interview sample to the wider population of lung cancer patients.

3.6 Conclusions

As only 20.5% of the sample reported pre-diagnosis Web searches, it seems like the role of the Web prior to a diagnosis of lung cancer is at present still limited, but this proportion is likely to increase in future, when barriers such as unfamiliarity with technology decrease, and willingness to be informed about one’s own health is likely to increase.

Participants perceived an impact of their Web searches on symptom appraisal, the decision to present to health services, and on how they communicated with doctors and requested referrals to specialist care. This suggests using the Web prior to diagnosis may impact on processes in the appraisal, help-seeking, and diagnostic interval referred to in Walter et al.’s model (Walter et al., 2012), and thus on the length of time until a diagnosis is made. While a study with a larger sample size and a validated measure of the pathway to diagnosis will be required to assess the statistical association between Web usage pre-diagnosis and time to diagnosis, my study highlights potential mechanisms of how Web-based health information may influence pathways to diagnosis.
Thus, it should be possible to utilise this potential. The study indicates that people with lung cancer and proxies do access Web-based information to help them appraise their condition, and the proportion doing so is likely to increase. We therefore need to ensure that those seeking information online can access useful information. In the following chapter, I show how I drew on findings from this exploratory study to inform the development of a Web-based intervention to encourage early presentation of lung cancer symptoms.
Chapter 4: Development of a Web-based intervention to encourage early help-seeking in people with symptoms associated with lung cancer

4.1 Introduction

In chapters two and three, I present findings from a systematic review of the literature on Web use for symptom appraisal, and findings from an exploratory survey and interview study with recently diagnosed lung cancer patients. From this work, I concluded that there is potential for online information to influence the time to diagnosis, and that people with lung cancer and proxies do access this information source (at present only a small proportion, but barriers are likely to decrease in future). This shows that there is potential for the Web to be a useful information source in the time before diagnosis. Additionally, the Web offers features such as round-the-clock availability, anonymity, widespread access, and the potential to disseminate information to large audiences at low cost (Cline & Haynes, 2001; Ziebland et al., 2004). In order to harness this potential, I developed a Web-based intervention to encourage help-seeking for people with possible lung cancer symptoms. In this chapter, I will show how I incorporated findings from my previous study and systematic review into this intervention, and I present the first version of the intervention website as it was implemented in the feasibility study (the website was altered for the subsequent randomised controlled trials; these alterations are described in the respective chapters).

Several public health campaigns have been launched in recent years in the UK to promote help-seeking for lung cancer symptoms, e.g. ‘I’ll tackle it soon’, ‘3 Week Cough’, ‘Be Clear on Cancer’, and ‘Detect Cancer Early’ (Gordon et al., 2012; Ironmonger et al., 2014). These campaigns usually disseminate brief, simple messages mostly addressing the cough symptom. Although they tend to result in increases in knowledge levels, the symptom appraisal/help-seeking interval often remains the same (Tustin, 2012). People with relevant symptoms who are aware of the campaigns sometimes describe not perceiving the information as relevant to themselves (Caswell et al., 2017; McLachlan et al., 2015). Furthermore, due to the campaigns’ strong focus on the cough symptom, people often remain unaware of other warning signs of lung cancer such as back ache, weight loss or fatigue (Birt et al., 2014; Caswell et al., 2017; McLachlan et al., 2015). These observations were confirmed during my interview study with recently diagnosed lung cancer patients (Chapter 3). Walter et al. (2015) recommend that
initiatives for improving lung cancer awareness should focus on multiple symptoms rather than single symptoms, as single symptoms rarely predict lung cancer reliably.

Thus, other intervention forms are needed which are perceived as personally relevant by individuals with pertinent symptoms, and which provide detailed information on the varied symptom profile of lung cancer.

At present, Web pages about lung cancer symptoms typically present a list of possible symptoms and a brief description of risk factors (Figure 4.1 - 4.2). However, to my knowledge no study has to date assessed whether this is the ideal form of presenting lung cancer information in order to ensure prompt help-seeking for warning signs. Moreover, most lung cancer websites are written above the recommended reading level for general lay audiences (Weiss et al., 2016), suggesting considerable room for improvement.

To address the question of whether alterations in the presentation of online health information about lung cancer could increase medical help-seeking, I developed an online intervention designed to encourage early help-seeking based on psychological theory, and designed a study to evaluate its effectiveness. I begin by describing how I used findings from my systematic review and interview study coupled with insights from consulting a cancer Patient and Public Involvement group, to develop the initial concept for the intervention. I then show how I implemented this concept by developing a tailored, theory-based website.

4.2 Overview of the intervention development process

The MRC Guidance for Developing and Evaluating Complex Interventions proposes four elements involved in the development and evaluation of an intervention (Figure 4.3) (Craig et al., 2008). The development element involves defining the problem with reference to empirical evidence and theory, as well as identifying mechanisms that sustain the problem and how these could be tackled. The piloting and feasibility element involves work to ensure the intervention can be delivered as intended and to inform sample size calculations for the evaluation element, which aims to assess effectiveness. The final element involves implementation and dissemination of the results to make findings accessible to decision-makers. These four elements are interlinked, and it is expected that the process is iterative rather than linear, with oscillations between the four elements. This chapter will focus on the development of the
intervention, and the following chapters will address feasibility (Chapter 5) and evaluation (Chapter 6 and 7). Possibilities for implementation and dissemination will be discussed in Chapter 8 (Final discussion).

I used an iterative approach in developing the intervention, incorporating views and feedback from potential users of the intervention at each stage. I began by drawing on my interview study (Chapter 3) to obtain an idea of the specific needs of individuals trying to appraise lung cancer symptoms that could be addressed by an online intervention. Next, I drew on previous qualitative research to identify relevant theoretical approaches. Combining the insights from the interview study with theoretical approaches, I then created a first rough draft of a possible intervention. After developing initial ideas and a first rough draft, I conducted Patient and Public Involvement work to obtain potential users’ feedback on the planned intervention components and select those that would be most relevant and acceptable. Using this feedback, I refined the intervention draft and conducted a Think Aloud study to assess whether users considered the intervention acceptable and feasible. I then completed a feasibility study, and used users’ feedback (from a free-text comments box) to further adapt the intervention, which I then tested again in a Think Aloud study. Thus, I cycled through a series of development-evaluation-development cycles (Yardley, Morrison, Bradbury, & Muller, 2015), adapting the intervention iteratively based on user feedback.

Figure 4.1 A screenshot of the NHS website on lung cancer symptoms (NHS choices, n.d.).
Figure 4.2 A screenshot of the cancer.org website on lung cancer symptoms (Cancer.org, n.d.).

Figure 4.3 MRC Guidance for developing and evaluating complex interventions from Craig et al., 2008 (2008).

4.3 Preliminary work and preparations

As suggested by Yardley et al. (2015), I adopted elements of the person-based approach and used a combination of qualitative research and a set of guiding principles to develop an intervention that is acceptable and persuasive to users, paying attention to the particular context of the target population.

4.3.1 Drawing on qualitative research with users and previous literature

As a first step, I drew on the systematic review of the literature (Chapter 2) as well as user perspectives from my interview study with lung cancer patients and proxies, which focused on
use of the Web to appraise symptoms and perceived impacts (Chapter 3). To inform the development of the intervention, I examined two topics in the interview data. Firstly, I explored how participants described their search and information appraisal processes, in order to understand how websites can be improved to better suit users’ needs, and what difficulties users may experience that need to be addressed. Secondly, I explored users’ perceptions of previous lung cancer awareness campaigns, as this can help us understand how a Web-based approach could complement these and address their limitations. In the research team, we discussed findings from the systematic review and the interviews and their implications for the intervention, and brainstormed for initial ideas of how these could be addressed (Table 4.1).

*Table 4.1 Incorporation of findings from the interview study and systematic review into the development of the intervention.*

<table>
<thead>
<tr>
<th>Finding from interview study/systematic review</th>
<th>Implication for the intervention</th>
<th>How could this be addressed in the intervention?</th>
</tr>
</thead>
<tbody>
<tr>
<td>People are aware of previous lung cancer awareness campaigns, but mostly cough (not other symptoms), and not very detailed knowledge</td>
<td>We need to find a way to present information about other symptoms as well without overwhelming users</td>
<td>Elicit symptoms from individual users first, and provide more in-depth information on these (tailored information)</td>
</tr>
<tr>
<td>Some people aware of previous lung cancer awareness campaigns assume the information is not relevant to them</td>
<td>We need to make sure the information presented is perceived as personally relevant</td>
<td>Present individually tailored information</td>
</tr>
<tr>
<td>Family/friends play an important role in help-seeking, and are often the trigger for initial consultation</td>
<td>We need to enhance users’ belief that other significant people (family/friends) want them to seek medical advice</td>
<td>Emphasise message is endorsed by family/friends</td>
</tr>
<tr>
<td>Process of elimination: comparing symptoms against those listed and discarding conditions for which the match is low (likely to happen with lung cancer as people typically only display 1-3 symptoms)</td>
<td>Rather than present individuals with a list of symptoms, we should present them with specific details on their particular symptoms</td>
<td>Personalisation of symptom information, presenting detailed information on endorsed symptoms</td>
</tr>
<tr>
<td>People often use online health information to prepare for consultations</td>
<td>We need to provide users with information they can easily take to their next consultation</td>
<td>Printable personalised summary of symptoms, risk factors and recommendations</td>
</tr>
</tbody>
</table>
### Finding from interview study/systematic review

<table>
<thead>
<tr>
<th>Finding from interview study/systematic review</th>
<th>Implication for the intervention</th>
<th>How could this be addressed in the intervention?</th>
</tr>
</thead>
<tbody>
<tr>
<td>People use online health information to support claims to their GP that further investigation of their symptoms is warranted (^1)</td>
<td>We need to provide users with guidance on when symptoms warrant further investigation, and information that can help them communicate with health professionals.</td>
<td>Personalised information on NICE guidelines for suspected cancer referral</td>
</tr>
<tr>
<td>People tend to trust known UK organisations (^r)</td>
<td>We need to enhance trust by showing that our message is endorsed by NHS health professionals.</td>
<td>Emphasise message is endorsed by health professionals.</td>
</tr>
<tr>
<td></td>
<td>We need to emphasise collaboration with organisations people know</td>
<td>Mark all pages with the University of Manchester and Medical Research Council logo. Emphasise that NHS practitioners reviewed the intervention.</td>
</tr>
</tbody>
</table>

\(^1\) finding derived from the interview study

\(^r\) finding derived from the systematic review

---

The exploration of the review and interview findings enabled me to develop an initial set of ideas for relevant intervention components that can address user needs and perspectives. To flesh out these ideas, I next explored which psychological theory could be used to ensure targeting of relevant constructs to change behaviour.

#### 4.3.2 Selecting a theoretical approach

In order to identify a relevant theoretical approach to use for the intervention development, I synthesised previous qualitative research on barriers to help-seeking in lung cancer. First I derived a list of barriers to help-seeking from the literature, and then sorted these barriers into broader themes. The themes were discussed in the research team until consensus was reached, to ensure they captured the findings conceptually and to ensure findings were allocated to all appropriate themes (Table 4.2).
Table 4.2 Barriers to help-seeking in lung cancer identified in previous literature.

<table>
<thead>
<tr>
<th>Barrier</th>
<th>Theme</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptoms experienced as minor, mild, unspecific</td>
<td>Symptom perception + Knowledge</td>
<td>(Tod et al., 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Corner et al., 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(McLachlan et al., 2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Corner et al., 2005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Birt et al., 2014)</td>
</tr>
<tr>
<td>Symptoms experienced as relating to other co-morbid conditions</td>
<td>Symptom perception</td>
<td>(Tod et al., 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Corner et al., 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Tod &amp; Joanne, 2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(McLachlan et al., 2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Birt et al., 2014)</td>
</tr>
<tr>
<td>Symptoms not perceived as connected, individual health changes are interpreted separately</td>
<td>Symptom perception + Knowledge</td>
<td>(Corner et al., 2006)</td>
</tr>
<tr>
<td>Symptoms experienced as intermittent, infrequent or insidious</td>
<td>Symptom perception + Knowledge</td>
<td>(McLachlan et al., 2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Birt et al., 2014)</td>
</tr>
<tr>
<td>Knowledge and awareness of lung cancer: Thought symptoms would be more severe, poor knowledge of risk factors</td>
<td>Knowledge</td>
<td>(Corner et al., 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Tod et al., 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Corner et al., 2005)</td>
</tr>
<tr>
<td>Fear and fatalistic beliefs about lung cancer and treatability</td>
<td>Perceived consequences of help-seeking</td>
<td>(Tod et al., 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Tod &amp; Joanne, 2010)</td>
</tr>
<tr>
<td>Worry about wasting the doctors’ time</td>
<td>Perceived consequences of help-seeking, Social acceptance regarding help-seeking</td>
<td>(Tod et al., 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Tod &amp; Joanne, 2010)</td>
</tr>
</tbody>
</table>

112
<table>
<thead>
<tr>
<th>Barrier</th>
<th>Theme</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fear of blame and stigma</td>
<td>Social acceptance regarding help-seeking</td>
<td>(Tod et al., 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Corner et al., 2006)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Tod &amp; Joanne, 2010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Corner et al., 2005)</td>
</tr>
<tr>
<td>Culture: Great value placed on stoicism, media advice not to present to primary care unless severe</td>
<td>Social acceptance regarding help-seeking</td>
<td>(Tod et al., 2008)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Tod &amp; Joanne, 2010)</td>
</tr>
<tr>
<td>Dislike of attending doctor’s surgery</td>
<td>Attitude</td>
<td>(Birt et al., 2014)</td>
</tr>
<tr>
<td>Limited access to health care/availability of appointments</td>
<td>Perceived/actual control over help-seeking</td>
<td>(Birt et al., 2014)</td>
</tr>
<tr>
<td>Absence of advice by GPs on when to return</td>
<td>Social acceptance regarding help-seeking</td>
<td>(Birt et al., 2014)</td>
</tr>
</tbody>
</table>

There are a variety of different approaches to (health) behaviour change. The four main perspectives include (Ogden, 2012):

1. Learning and cognitive theories which focus on reinforcement and incentives
2. Stage models which identify different factors important for behaviour change based on the stage of ‘readiness to change’ an individual is in
3. Affect-based approaches which recognise that people do not always process information rationally and account for affect-driven behaviour
4. Social cognition theories that predict behaviour based on perceived consequences and perceived social acceptance of the behaviour.

As Table 4.2 shows, aside from knowledge and perceptions of symptoms, perceived consequences, attitudes, social acceptance and perceived control seem to play important roles. This suggests that social cognition theories are of particular relevance.

Help-seeking behaviour is characterised as a “problem focused, intentional action and interpersonal interaction” (Cornally & McCarthy, 2011, p. 282). One of the most prominent social cognition theories, the Theory of Planned Behaviour (TPB), explains behaviour as a planned
action, influenced mainly through an individual's intentions, which in turn are shaped by beliefs around social norms, outcomes of behaviour, and control over the behaviour (Ajzen, 1985). Thus, this theory seems of particular relevance when studying help-seeking behaviour and captures the themes identified in Table 4.2.

I decided to focus on TPB based on its relevance to the barriers to help-seeking, as well as the following reasons:

a. TPB has been successfully used to predict help-seeking for cancer symptoms (Hunter, Grunfeld, & Ramirez, 2003)
b. a meta-analysis on Internet-based behaviour change interventions found larger effects for TPB than other theoretical approaches (Webb, Joseph, Yardley, & Michie, 2010),
c. TPB has been successfully used to inform the development of a previous (non Web-based) intervention to target help-seeking in lung cancer (S. M. Smith et al., 2012),
d. TPB takes an individual’s perception of their social environment, consequences of the behaviour, and control over the behaviour into account, and previous literature shows these are key factors influencing lung cancer patients' decisions to seek medical help (Birt et al., 2014; Corner et al., 2005, 2006; McLachlan et al., 2015; Tod et al., 2008; Tod & Joanne, 2010).

The TPB postulates that behaviour is a function of intention to perform the behaviour, which is influenced by beliefs: Beliefs about the behaviour and its outcomes, normative beliefs about social expectations, and beliefs about control over executing the behaviour. Beliefs about behavioural outcomes are weighted by the evaluation of these outcomes, and normative beliefs are weighted by the motivation to comply with perceived social expectations. Beliefs about factors limiting or facilitating control over the behaviour are weighted by the perceived power of these factors. These weighted beliefs in turn influence attitudes towards the behaviour, perceived social norms, and perceived behavioural control (Figure 4.4). Perceived behavioural control can also have direct effects on the behaviour.
4.3.3 Combining findings from interviews with theory

Based on the exploration of the interview findings as well as the assessment of relevant psychological theory, I concluded that the intervention should include two central factors: 1) tailoring of information in order to enhance perceived personal relevance, to mitigate the ‘process of elimination’ and to present information on multiple symptoms (not only cough) without overwhelming users and 2) incorporation of components to address key beliefs which may impact on execution of behaviour.

4.3.4 Patient and public involvement

Once these initial ideas were conceptualised, I discussed these ideas and proposed content with members of a Patient and Public Involvement (PPI) group consisting of people who have been affected by cancer.

Before the meeting, the group members were sent a PowerPoint file that contained a rough draft of what the intervention could look like. The group members were informed about the rationale and aim of the intervention, and were requested to consider the following questions:

- How can we tailor the information to individual needs?
- How can we best influence users’ beliefs?
- Is the proposed information clear and easy to understand?
  - If not, how could this be clarified?
• Is the proposed phrasing acceptable?
  o If not, how could it be phrased more clearly?
• How can we make the website appear credible?

During the meeting (four people attended), I discussed these questions with the group members in detail, and noted any suggestions made by group members for changes, additions or removals of content. To decide whether to “accept” or “reject” user preferences during usability and PPI work, Curtis et al. (2015) recommend incorporating user suggestions which are (1) in line with the target behaviour, (2) compatible with the theoretical basis of the intervention, (3) compatible with usability recommendations and (4) easy to implement online. We used these criteria and thus recommendations made by the group were incorporated into the design of the intervention if they aimed to enhance help-seeking behaviour, were in line with the TPB, were compatible with usability recommendations such as those recommended by the National Institute on Aging for senior-friendly websites (National Institute on Aging, n.d.), and if they were feasible to implement online. The main points derived from the PPI work are shown in Box 1.

One of the key outcomes of this work was the group’s suggestion that quotes (accompanied by names and roles) may be a suitable method of presenting persuasive information. This fed into the operationalisation of the initial ideas into a concrete website, which I describe in detail in the next section.

4.3.5 Guiding principles

According to Yardley et al. (2015), it is useful to produce a set of guiding principles which can be consulted throughout planning and development of an intervention and which ensure it relates to its key objectives. My guiding principles were informed by the interview study (Mueller, Jay, Harper, & Todd, 2017), the systematic review (Mueller, Jay, Harper, Davies, et al., 2017) as well as previous research on help-seeking in lung cancer (Birt et al., 2014; Corner et al., 2005, 2006; McLachlan et al., 2015; Tod et al., 2008; Tod & Joanne, 2010) (Box 2). These guiding principles were referred to throughout the intervention development to provide a coherent focus.
Once I had created a first draft of the intervention design together with the PPI group and established guiding principles, I proceeded to develop the intervention website. In the following section, I describe how information was tailored, and how I used theory to inform the intervention development.

Box 1. Suggestions made by the Patient and Public Involvement group/thoughts voiced by the group regarding our initial ideas.

- The group suggested that some form of source credibility should be added to enhance impact on peoples’ beliefs, i.e. endorsement by health professionals or other relevant people. The group therefore suggested including quotes by health professionals, accompanied by names and affiliations, to emphasise key messages.
- Credibility: Credibility is very important and endorsements by health professionals might be beneficial. A specific page should be dedicated to providing information about the research team, where the medical information and recommendations were derived from, who reviewed the website etc.
- Avoiding the term cancer: The group pointed out that the term ‘lung cancer’ should not be sprung upon users too quickly, as it could cause anxiety and distress. The group suggested putting stronger focus on the symptoms and when to seek help, and avoiding the term cancer wherever possible.
- Avoiding self-diagnosis: The group pointed out that users might use the intervention as a self-diagnosis tool. They therefore suggested emphasising that the website is designed to provide advice on when to seek help, rather than self-diagnosis.
- Medical disclaimer: The group recommended emphasising to users that the website should not be used to replace advice from a medical professional.
- There was some confusion regarding terminology used to refer to health professionals. The group recommended using the term ‘doctor’ consistently, possibly with General Practitioner/GP in brackets.
4.4 Incorporating TPB into the intervention

Interventions designed to change or elicit behaviour can target one or more of the determinants of behaviour theorised in the TPB; this can include replacing existing beliefs as well as introducing new beliefs (Ajzen, 2006a). In order to target TPB constructs, I proceeded in three steps:

1) Identifying relevant beliefs that play a role in help-seeking in lung cancer and linking these to TPB constructs

2) Identifying behaviour change techniques (BCTs) that can target these TPB constructs and beliefs

3) Operationalising and implementing the identified BCTs in a Web-based intervention

4.4.1 Identification of relevant beliefs and linking these to TPB constructs

I used findings from previous studies (Table 4.2) as well as the interview findings (Chapter 3) to develop a list of beliefs that appear to play a role in help-seeking for lung cancer symptoms. Next, these beliefs were mapped onto the three types of beliefs identified in the TPB (beliefs about outcomes of the behaviour, normative beliefs, and control beliefs) by using a consensus-based approach in the research team (Table 4.3).
Box 2. Guiding principles.

1. **Intervention objectives**
   - To provide information to people with symptoms potentially related to lung cancer, and encourage medical help-seeking where appropriate.

2. **Relevant aspects of users in the context**
   - People with lung cancer are often:
     - older
     - less technically literate
     - of lower socioeconomic/educational level
   - Those looking up symptoms may be family members/friends rather than afflicted individuals themselves

3. **Key behavioural issues, needs, or challenges that the intervention needs to address**
   - Barriers to help-seeking identified in the literature and in the interview study:
     - Symptoms are experienced as too mild to warrant help-seeking
     - Symptoms are perceived as related to existing conditions
     - People think symptoms of lung cancer would be more severe, poor knowledge of risk factors
     - Fear and fatalistic beliefs about lung cancer and treatability
     - Fear of wasting the doctor’s time/ being seen as a time-waster
     - Fear of blame and stigma (due to smoking)
     - Culture: Great value placed on stoicism, media advice not to present to primary care unless severe
     - Dislike of attending doctor’s surgery
     - Difficulties accessing healthcare (e.g. competing responsibilities, limited availability of appointments)
     - Information (e.g. from media campaigns) is perceived as irrelevant for the individuals’ situation

4. **Intervention design features that can address the barriers and achieve the aim**
   - Tailoring
     - This will ensure users are informed about their specific symptoms and risk factors, and when help-seeking is warranted
     - It will also make the information appear more personally relevant
   - Theory-based components
     - Will address beliefs such as fear, fatalistic beliefs, fear of wasting the doctors’ time, fear of social stigma
   - Provide information on when a referral for a chest x-ray is warranted
     - In order to help individuals communicate with their doctors during the diagnostic process (and enhance tailoring)
   - The intervention should be kept as close as possible to the usual experience of using the Web to appraise symptoms (as we are targeting people engaging in this activity), i.e. brief, once-off visit to the website, no multiple visits required
Table 4.3 Beliefs about help-seeking identified from the literature, mapped onto TPB constructs.

<table>
<thead>
<tr>
<th>Beliefs about outcomes</th>
<th>Normative beliefs</th>
<th>Control beliefs</th>
</tr>
</thead>
<tbody>
<tr>
<td>If medical advice is sought, no serious cause will be found.* (Birt et al., 2014;</td>
<td>Worry about being seen as a time-waster by doctors. * (Tod &amp; Joanne, 2010)</td>
<td>Perceived difficulties due to limited access to health care / availability of appointment* (Birt et al., 2014)</td>
</tr>
<tr>
<td>Corner et al., 2005, 2006; McLachlan et al., 2015; Tod et al., 2008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fear and fatalistic beliefs about lung cancer and treatability (seeking medical</td>
<td>Culture: Great value placed on stoicism, media advice not to present to primary</td>
<td></td>
</tr>
<tr>
<td>advice might be pointless if it is lung cancer, because lung cancer cannot be</td>
<td>care unless severe (Tod et al., 2008) (Tod &amp; Joanne, 2010)</td>
<td></td>
</tr>
<tr>
<td>treated) (Tod et al., 2008; Tod &amp; Joanne, 2010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worry about wasting the doctors' time* (Tod et al., 2008)</td>
<td>Fear of blame and stigma (due to smoking) (Corner et al., 2005, 2006; Tod et al.,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2008; Tod et al., 2008; Tod &amp; Joanne, 2010)</td>
<td></td>
</tr>
<tr>
<td>* Also observed during the interview study</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.4.2 Identification of behaviour change techniques

In the previous step, I identified the beliefs that needed to change according to the TPB in order to impact on help-seeking. The next logical step was to consider which techniques to use in order to change these beliefs.

Michie et al. (2008) developed a comprehensive list of behaviour change techniques (BCTs) that are used in behaviour-change interventions, based on a systematic review of the literature and relevant textbooks, as well as “brainstorming” among behaviour-change experts. BCTs are defined as observable mechanisms of change or ‘active ingredients’ used in behaviour change interventions (Michie et al., 2013).

Michie et al. (2008) also created a matrix which maps BCTs onto theoretical construct domains found in a wide array of psychological behaviour-change theories. I used this matrix to identify which BCTs to use in my intervention. I first needed to identify which of the construct domains in Michie et al.’s (2008) matrix most closely matched the TPB constructs I wanted to target. Two independent researchers undertook this matching exercise, achieving 100% agreement: ‘Beliefs about consequences’ were matched with ‘Behavioural beliefs’, ‘Social influences’ were matched with ‘Normative beliefs’, and ‘Beliefs about capabilities’ were matched with ‘Control beliefs’. We
then collated a list of all BCTs considered to be effective in changing these constructs according to (Michie et al., 2008).

Once I had identified a list of possible BCTs to target the TPB constructs, I assessed these BCTs for their suitability for the intervention based on the following criteria:

- Amenability to mode of delivery (Web-based, one-off short-term interaction of about 20 minutes; users will not return to the website at later time),
- Relevance and suitability for the targeted behaviour: Making an appointment with the GP to have the symptoms checked
- Relevance and suitability to target the beliefs identified in Table 4.2.

Table 4.4 shows the list of BCTs I collated, and the assessment of their suitability for the present context. Following the assessment of the suitability of BCTs, I identified the BCTs shown in Table 4.5 as suitable and likely to be effective in the intervention.

4.4.3 Operationalising behaviour change techniques

The identified BCTs now needed to be operationalised for this particular context. As mentioned above, the PPI group suggested I use quotes to substantiate the messages. I decided to use this medium to convey information to target beliefs. To illustrate, I now show examples below of how TPB constructs were targeted using the BCTs.

**Behavioural beliefs**

Beliefs about the outcomes of help-seeking were targeted by showing quotes from health professionals which emphasise positive outcomes of help-seeking (Figure 4.5), thus using “Information regarding behaviour, outcome”. Furthermore, the quote is endorsed by a credible source and presents a pro-argument for early presentation (improved treatability), and thus constitutes the BCT “persuasive communication”.

Table 4.4 Construct domains and relevant behaviour change techniques identified from (Michie et al., 2008).

<table>
<thead>
<tr>
<th>Construct domain</th>
<th>Maps on to TPB construct</th>
<th>Behaviour change techniques</th>
<th>Suitability</th>
</tr>
</thead>
</table>
| Beliefs about consequences | Behavioural beliefs | Self-monitoring                                               | Not suitable  
This would have to involve monitoring over a longer period of time – not amenable to mode of delivery |
|                      |                          | Information regarding behaviour, outcome;                   | Suitable                                                                   |
|                      |                          | Persuasive communication                                    | Suitable                                                                   |
|                      |                          | Feedback                                                     | Not suitable  
This would have to involve monitoring over a longer period of time – not amenable to mode of delivery |
| Social influences   | Normative beliefs        | Social processes of encouragement, pressure, support         | Suitable                                                                   |
|                      |                          | Modelling/demonstration of behaviour by others               | Not suitable  
Not suitable for the targeted behaviour; difficult to ‘show’ people making an appointment on a website; and probably not necessary as the physical act of making an appointment does not pose difficulties. |
| Beliefs about capabilities | Control beliefs     | Increasing skills: problem-solving, decision-making, goal-setting | Suitable                                                                   |
|                      |                          | Self-monitoring                                              | Not suitable  
This would involve monitoring over a longer period of time – not amenable to mode of delivery |
<table>
<thead>
<tr>
<th>Construct domain</th>
<th>Maps on to TPB construct</th>
<th>Behaviour change techniques</th>
<th>Suitability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Graded task, starting with easy task</td>
<td>Not suitable</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>This would involve monitoring over a longer period of time – not amenable to mode of delivery. Also not suitable for the targeted behaviour, as it is difficult to design ‘graded tasks’ leading up to making an appointment.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coping skills</td>
<td>Not suitable</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Not suitable for the targeted behaviour, which only involves a phone call to make an appointment.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rehearsal of relevant skills</td>
<td>Not suitable</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>This would have to involve monitoring over a longer period of time – not amenable to mode of delivery. Also not suitable for the targeted behaviour, which only involves a phone call – rehearsal of skills not needed.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Social processes of encouragement, pressure, support</td>
<td>Suitable but in this context not suitable to target control beliefs; more suitable to target normative beliefs. Based on findings from the literature and the interview study, people do not lack the belief that they are able to complete the behaviour (making an appointment); rather they lack the belief that they can fit this in with existing responsibilities or appointment availability at their clinic. Social encouragement is unlikely to target these beliefs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feedback</td>
<td>Not suitable</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>This would have to involve monitoring over a longer period of time – not amenable to mode of delivery</td>
</tr>
</tbody>
</table>
Table 4.5 Behaviour change techniques (BCTs) identified as suitable and likely to be effective in the TPB-based intervention.

<table>
<thead>
<tr>
<th>Construct domain</th>
<th>Maps on to TPB construct</th>
<th>BCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beliefs about consequences</td>
<td>Behavioural beliefs</td>
<td>Information regarding behaviour, outcome; Persuasive communication</td>
</tr>
<tr>
<td>Social influences</td>
<td>Normative beliefs</td>
<td>Social processes of encouragement, pressure, support</td>
</tr>
<tr>
<td>Beliefs about capabilities</td>
<td>Control beliefs</td>
<td>Increasing skills: problem-solving, decision-making, goal-setting</td>
</tr>
</tbody>
</table>

Figure 4.5 Screenshot showing how I targeted behavioural beliefs on the intervention website (version 1).

Normative beliefs

Based on the literature, I identified two groups of people whose acceptance or sanctioning of help-seeking seems to shape individuals’ normative beliefs: Significant others (e.g. family and friends) and health professionals. Users were presented with quotes from (fictional) family members and health professionals, stating that they endorse early help-seeking (Figures 4.6 and 4.7). These quotes make use of the BCT “Social processes of encouragement, pressure, support”.

---

“If lung cancer is caught early, we can do more to treat it. So it’s important to see your doctor with any symptoms that persist for 3 weeks or longer.”

— Yvonne Summers, Consultant Medical Oncologist at the Christie Hospital, Manchester
Figure 4.6 Screenshot showing how I targeted normative beliefs about family/friends on the intervention website (version 1).

Figure 4.7 Screenshot showing how I targeted normative beliefs about health professionals on the intervention website (version 1).

Control beliefs

To enhance perceived behavioural control over help-seeking, a step-by-step guide to making an appointment was provided (Figure 4.8). This TPB-component aimed to enable users to set specific goals, and enhance their confidence in achieving these goals by highlighting helpful resources (e.g. a website on how to register with a GP). Thus, this component made use of the BCT “Increasing skills: problem-solving, decision-making, goal-setting”.

As discussed above, the intervention included two key factors: the incorporation of theory-based components (described in the preceding section), and information tailoring, which I will now describe in detail.
4.5 Tailoring

The interview study suggested that information from awareness raising campaigns is less likely to lead to help-seeking when individuals are optimistically biased and assume the information does not apply to them personally, which is confirmed in previous research (Caswell et al., 2017; McLachlan et al., 2015). Research has indicated that tailored information, as compared to general information, can enhance perceived personal relevance (Kreuter, Farrell, Olevitch, & Brennan, 2012), and can increase detection of cancer symptoms and intentions to seek help (de Nooijer, Lechner, Candel, & de Vries, 2004). Thus, my website provided users with individually tailored information about symptoms, risk factors and advice on whether medical help should be sought. On entry to the website, users first completed a set of questions designed to elicit data which were then used for information tailoring.

4.5.1 Symptoms and risk factors

Users were asked to report which symptoms they experienced, their age, and smoking status. For each endorsed symptom, users were asked how long it had persisted (<3 weeks or ≥3 weeks) and whether they experienced it as very intense/severe. The website then presented
detailed information to participants on symptoms and risk factors they endorsed (Figure 4.9 and 4.10).

Additionally, the intervention outputted a summary listing all risk factors and symptoms a user reported, which could be printed or saved as pdf. The summary also provided tailored information on whether a referral for an urgent chest X-ray might be indicated based on NICE clinical guidelines (National Institute for Health and Care Excellence, 2015) (Figure 4.11).

![Information about your symptoms](image1)

**Cough**

**What to look for:**
A cough in lung cancer can take different forms. It can be dry and tickly, or productive (with phlegm). It can be relatively mild and may even improve at times.

**What to do:**
If you have been coughing for 3 weeks or longer, it is recommended that you report your cough to a doctor even if you think you know the cause of the cough, or if it is just a slight cough if your cough is severe, or if you feel concerned for any other reason, you may want to contact your doctor sooner.

A persistent cough, even if you are a smoker, is telling you something is wrong and needs to be examined.

**Unintentional weight loss**

**What to look for:**
Your body weight may rise and fall regularly, but if you unintentionally lose more than 5% of your usual weight over 6 - 12 months, this may be a cause for concern. You should pay particular attention if you notice other symptoms, such as tiredness, loss of appetite or a change in toilet habits.

**What to do:**
Our weight calculator calculated a weight loss of 7%. It is recommended that you report this to a doctor.

![Figure 4.9](image2)

The screenshot shows an example of tailored symptom information a user would receive who endorsed a cough and unintended weight loss during the initial survey (website version 1).

![Your risk factors](image3)

**Are ex-smokers at risk?**

- Smoking is the biggest risk factor for lung cancer
- Tobacco contains toxic substances which are known to be cancer-producing
- Includes cigarette smoking and other tobacco products such as cigars, pipe tobacco, snuff and chewing tobacco
- While quitting smoking greatly reduces your risk, your risk will still be higher than that of someone who has never smoked
- After about 10 years of stopping, the risk of an ex-smoker will be about half that of a smoker

**What to do:** If you are concerned about your symptoms, mention your smoking history, even if you have stopped smoking several years ago. This information is important for your doctor to assess your risk and to make sure all important diagnostic tests are done. You will not receive any different treatment or health care if you are an ex-smoker.

**Age as a risk factor**

- Older age is a risk factor for developing lung cancer
- Lung cancer typically develops in people above the age of 50
- The average age when people are diagnosed is 70

![Figure 4.10](image4)

The screenshot shows an example of tailored risk factor information a user would receive who reported being an ex-smoker, and age above 40 (website version 1).
4.5.2 TPB constructs

On entry to the intervention website, participants were presented with a list of statements representing the TPB constructs I targeted (Table 4.6), and asked “Which of the following statements do you agree with most?” This question aimed to identify which beliefs were most likely to impede an individual’s help-seeking behaviour.

Table 4.6 Statements used to tailor TPB components to individual users.

<table>
<thead>
<tr>
<th>Statement</th>
<th>TPB construct</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Making an appointment to see a doctor about these symptoms would be pointless/harmful/bad.</td>
<td>Behavioural beliefs</td>
</tr>
<tr>
<td>b) It is important whether others (family and friends) think it is necessary to see a doctor about these symptoms.</td>
<td>Normative beliefs – family/friends</td>
</tr>
<tr>
<td>c) It is important whether doctors think the symptoms are worth investigating.</td>
<td>Normative beliefs – health professional</td>
</tr>
<tr>
<td>d) Getting an appointment to report these symptoms to a doctor would be difficult.</td>
<td>Control beliefs</td>
</tr>
</tbody>
</table>

Based on their selection, participants received tailored information to target the TPB construct they chose. If a user selected statement c, for example, the user would be presented with the quote shown in Figure 4.7 to target normative beliefs about health professionals.
4.5.3 Affected individual or proxy?

I provided information both for users who were researching their own symptoms and those searching on behalf of someone else (“proxy”). While all users received the same information, the wording was adjusted for proxies, e.g. “If you have been coughing for 3 weeks or longer, it is recommended that you report your cough to a GP” was adjusted for proxies to “If your friend/relative has been coughing for 3 weeks or longer, it is recommended that they report their cough to a GP.”

4.6 Medical content

The medical content of the intervention is based on information from:

1) Medical reference books on respiratory medicine (Bourke & Burns, 2011; Dilworth & Baldwin, 2001; Howard, 1991; Longmore et al., 2014; Maskell & Millar, 2009)

2) Existing websites from UK cancer charities as well as the NHS (Table 4.7)

3) Peer-reviewed articles (Hamilton, Peters, Round, & Sharp, 2005; McWilliams et al., 2013; Wao, Mhaskar, Kumar, Miladinovic, & Djulbegovic, 2013)

4) Guidelines for lung cancer treatment developed by the UK National Collaborating Centre for Cancer (National Collaborating Centre for Cancer (UK), 2011)


Table 4.7 Details of Web pages used to inform the intervention content.

<table>
<thead>
<tr>
<th>Charity / Organisation</th>
<th>Title of the page</th>
<th>Link (All websites accessed and archived on 20 September 2016)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foundation (the Roy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Castle Lung Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foundation, n.d.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NHS Choices (NHS</td>
<td>Lung cancer myths and facts</td>
<td><a href="http://www.nhs.uk/Livewell/Lungcancer/Pages/Lungcancermythsandfacts.aspx">http://www.nhs.uk/Livewell/Lungcancer/Pages/Lungcancermythsandfacts.aspx</a></td>
</tr>
<tr>
<td>Choices, n.d.-h)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Charity / Organisation</td>
<td>Title of the page</td>
<td>Link (All websites accessed and archived on 20 September 2016)</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Patient.co.uk (Patient.co.uk, n.d.-b)</td>
<td>Hoarseness</td>
<td><a href="http://www.patient.co.uk/doctor/hoarseness-pro">http://www.patient.co.uk/doctor/hoarseness-pro</a></td>
</tr>
<tr>
<td>NHS Choices (NHS Choices, n.d.-e)</td>
<td>Lumps and swellings</td>
<td><a href="http://www.nhs.uk/conditions/lumps-swellings/Pages/Introduction.aspx#face">http://www.nhs.uk/conditions/lumps-swellings/Pages/Introduction.aspx#face</a></td>
</tr>
<tr>
<td>Patient.co.uk (Patient.co.uk, n.d.-a)</td>
<td>Bronchiectasis</td>
<td><a href="http://www.patient.co.uk/health/bronchiectasis-leaflet">http://www.patient.co.uk/health/bronchiectasis-leaflet</a></td>
</tr>
<tr>
<td>NHS Choices</td>
<td>Lung cancer: Are you at risk?</td>
<td><a href="http://www.nhs.uk/Livewell/Lungcancer/Pages/Lungcancerareyouatrisk.aspx">http://www.nhs.uk/Livewell/Lungcancer/Pages/Lungcancerareyouatrisk.aspx</a></td>
</tr>
</tbody>
</table>
The website was developed in close collaboration with a software engineer with Web development experience, who is also a trained cardiac nurse and thus has medical domain knowledge.

Website content was reviewed for medical accuracy and appropriateness by an NHS consultant medical oncologist specialising in lung cancer, an NHS General Practitioner, and a trained nurse. Where these individuals voiced concerns, the content was adapted accordingly. I dedicated a specific page on the website, ‘About us’, to providing information about the research team, the sources used to inform the medical content, and the individuals who reviewed the information. Written content was subjected to a Flesch–Kincaid reading level test, to ensure texts were close to the recommended reading level for the general public (grade 8) (Eltorai et al., 2014).

One should also bear in mind that my intervention was not designed to promote self-diagnosis; the focus was on identifying when symptoms should be presented to health services. Users were informed of this distinction on the study homepage.

4.7 Think Aloud evaluation

Once a first version of the website was completed, I conducted a Think Aloud evaluation to assess whether the website was easy to navigate and use, and whether it was acceptable and engaging. In the Think Aloud paradigm, users navigate a website while voicing their thoughts about it and vocalising their actions, thus helping the researcher to identify useful, less useful and confusing features (Krug, 2013; Lewis, 1982). Ethical approval was granted by the School of Computer Science Ethics Committee (approval number: CS 206).

4.7.1 Participants

Five users participated in this evaluation, aged 22-55 years (3 female, 2 male).

4.7.2 Procedure

Participants were presented with symptom vignettes and asked to imagine they (or someone in their family) were experiencing the symptoms in the vignette (Box 3). Users were instructed to click through the website and use the information to appraise the symptoms described in the
vignette while verbalising their thoughts. Verbal prompts were used to elicit further information regarding the following questions:

- Are the content and the intervention components acceptable, feasible and salient?
- Are the questions and the text clear and easy to understand?
- Are all instructions clear and easy to follow?
- How long did it take users to click through the website?
- Is the length of time acceptable?
- Do users find the completion of the battery of measures after the intervention acceptable?

**Box 3. Example of a symptom vignette shown to users during the usability evaluation.**

A family member of yours is 56 years old and has been smoking for the past 20 years. She’s always had a smoker’s cough, but lately it has become a little worse and sounds ‘barking’. You have tried to persuade her to go see a doctor, but she refuses because she says you should not go to a GP with a mere cough. She insists it is only a smoker’s cough and she doesn’t want to waste the doctor’s time. You first noticed these changes about a month ago.

### 4.7.3 Analysis

While participants completed the intervention, I took notes on any comments made, and on any pages/elements during which participants appeared confused or hesitant. I did not answer questions or help participants with the completion, to emulate the situation of actual website users. After all five users had completed the evaluation, I assessed comments and suggestions against criteria developed by Curtis et al. (2015), and implemented user suggestions if they were (1) relevant to the target behaviour, (2) available online, (3) sufficiently easy to implement, (4) aligned with usability and user experience recommendations, and (5) supported from theoretical findings and/or evidence. Suggestions were also assessed against the guiding principles (Box 2).
4.7.4 Results

The website took 11-17 minutes (average 14 minutes) to complete, which users deemed acceptable. As Table 4.8 shows, feedback from the Think Aloud evaluation prompted changes to the phrasing of measures and intervention content, the structure, visual design, and information regarding credibility. Overall, the questions and information provided were deemed acceptable (with some minor changes to phrasing, see below). Participants particularly commented positively on quotes endorsed by health professionals. User behaviour and comments suggested that information about risk factors was likely to be skimmed or skipped.

Table 4.8 Summary of feedback received and observations made during the Think Aloud evaluation.

<table>
<thead>
<tr>
<th>Problems</th>
<th>How addressed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acceptability and salience of information</strong></td>
<td></td>
</tr>
<tr>
<td>Would just skim over TPB quotes, though when realised it was quotes by a consultant, paid more attention</td>
<td>Name and role of the person belonging to the quote was highlighted to stand out</td>
</tr>
<tr>
<td>The summary is helpful because it provides details on the symptoms, their context, and the outcome (advice)</td>
<td>No changes required.</td>
</tr>
<tr>
<td>The symptom information seems important so I would not skim this but read it properly</td>
<td>No changes required.</td>
</tr>
<tr>
<td>Would just skim the risk factor information, as I feel I already know this</td>
<td>No changes required, but interesting to note.</td>
</tr>
<tr>
<td><strong>Credibility</strong></td>
<td></td>
</tr>
<tr>
<td>Was the information checked by any health professionals, or just researchers from the University? If the former, it should say that.</td>
<td>This information was added to the ‘About us’ page.</td>
</tr>
<tr>
<td>It’s not very clear that it is a UK website and based on NICE guidelines.</td>
<td>This was emphasised on the study homepage.</td>
</tr>
<tr>
<td><strong>Phrasing</strong></td>
<td></td>
</tr>
<tr>
<td>‘Did the [symptom] come on suddenly or is it very severe?’ the ‘come on suddenly’ might be difficult to answer, and difficult to distinguish from ‘severe’</td>
<td>This was changed just to ‘Is the [symptom] very severe?’ The point of this question was to identify people who have not had their symptom for 3 weeks but whose symptom should still be presented due to urgency. Removing the first part of the question did not change this and made it clearer/less confusing.</td>
</tr>
<tr>
<td>In the question on tiredness, you should add ‘Have you felt tired <em>for unexplained reasons</em>?’</td>
<td>This was added, as participants may be more tired for normal reasons such as lack of sleep.</td>
</tr>
<tr>
<td>Problems</td>
<td>How addressed</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>There should be more options to answer ‘not sure’ on the page with</td>
<td>Options were added where symptoms are not easily visible to proxies, e.g. haemoptysis.</td>
</tr>
<tr>
<td>questions on symptoms, for users filling the form in on someone else’s</td>
<td></td>
</tr>
<tr>
<td>behalf</td>
<td></td>
</tr>
<tr>
<td>In the question ‘Have you experienced a change in a long-standing cough?’</td>
<td>‘long standing’ was highlighted in bold.</td>
</tr>
<tr>
<td>the ‘long-standing’ should be emphasised as it is otherwise easily</td>
<td></td>
</tr>
<tr>
<td>missed</td>
<td></td>
</tr>
<tr>
<td>When asking participants whether they would like to complete the</td>
<td>This was changed to “I’d like to help”, as suggested by the participant.</td>
</tr>
<tr>
<td>optional questionnaire at the end, the option “Sure, I’ll help” sounds</td>
<td></td>
</tr>
<tr>
<td>too informal; might be off-putting for older users</td>
<td></td>
</tr>
<tr>
<td>Website structure</td>
<td></td>
</tr>
<tr>
<td>Perhaps the ‘Print summary’ option should be at the very end</td>
<td>This option is on the last of the information pages, thereafter only questionnaire pages follow. We were unable to add this to the very last page, as this would be the optional questionnaire.</td>
</tr>
<tr>
<td>The end of the study is quite abrupt, it’s unclear when it’s finished</td>
<td>We added a message that appears at the end, telling participants that they have now completed the study and will be redirected.</td>
</tr>
<tr>
<td>The final questionnaire is too long, and the questions seem redundant.</td>
<td>We shortened the TPB questionnaire to one item per construct.</td>
</tr>
<tr>
<td>Visual design</td>
<td></td>
</tr>
<tr>
<td>The quote under the image of the doctor is too close, there should be</td>
<td>The space was increased.</td>
</tr>
<tr>
<td>more space</td>
<td></td>
</tr>
<tr>
<td>The notification for missed question works well</td>
<td>No changes required.</td>
</tr>
<tr>
<td>The 2nd page of the final questionnaire looks the same as the first,</td>
<td>We added a banner to the top which states ‘Page 1’ and ‘Page 2’.</td>
</tr>
<tr>
<td>which might be confusing.</td>
<td></td>
</tr>
<tr>
<td>Minor grammar/spelling/oversights</td>
<td></td>
</tr>
<tr>
<td>In a few places we had not adjusted the wording to proxies</td>
<td>Wording was adjusted appropriately.</td>
</tr>
<tr>
<td>‘Have you experienced any expected weight loss?’ This should say</td>
<td>Changed to unexpected weight loss.</td>
</tr>
<tr>
<td>unexpected weight loss</td>
<td></td>
</tr>
</tbody>
</table>

**4.7.5 Conclusions**

The Think Aloud evaluation mainly helped us to identify issues with the phrasing of intervention content and measures that were confusing, unclear or potentially off-putting. Solutions were put forward by the participants, which helped us to resolve these problems in a suitable manner.

The evaluation also allowed us to identify what material is likely to be read in detail, and what is
more likely to be skimmed. Importantly, the evaluation also helped us to understand how users assess the credibility of the website. By following users’ suggestions, I was able to add information and features that potentially help to enhance the perceived credibility. Finally, the Think Aloud evaluation helped to improve the visual and structural design of the Web application.

4.8 Summary

To summarise, I developed a tailored, theory-based intervention to encourage early help-seeking for symptoms potentially related to lung cancer. The intervention is tailored to users’ individual symptoms and risk factors, and provides tailored advice on whether medical help should be sought. I developed this intervention using findings from previous literature on help-seeking in lung cancer, my own interview study findings, medical reference works, and TPB literature. Development of this intervention was also informed by discussions with members of the public, and a Think Aloud study using fictional symptom scenarios. I used a combination of qualitative research with potential users and a clear set of guiding principles to ensure the intervention targeted its key objectives and addressed user needs and preferences (Yardley et al., 2015).

An overview of the structure of the intervention is provided in Figure 4.12. I aimed to test the effectiveness of this intervention in an online randomised trial. Prior to implementation of the full trial, feasibility work was necessary to determine whether I would be able to reach the target population, whether the intervention components worked together smoothly, and to inform sample size calculations (Arain, Campbell, Cooper, & Lancaster, 2010; Craig et al., 2008). In the following chapter, I will describe the methods, design and findings of the feasibility study.
Users complete an initial survey on their symptoms, risk factors and TPB constructs

Users receive tailored feedback about symptoms

Users receive tailored feedback about risk factors

Users receive tailored TPB components to target specific beliefs about help-seeking

Users receive a tailored summary, including information on whether an urgent chest x-ray might be indicated based on NICE clinical guidelines

*Figure 4.12 Structure of the version of the intervention implemented in the feasibility study.*
Chapter 5: Feasibility study

5.1 Aim

I conducted a feasibility study for a randomised controlled trial to test the intervention described in the previous chapter. My main aims for this study were (i) to determine whether it would be possible to recruit the target population of people with undiagnosed symptoms potentially related to lung cancer, (ii) to obtain estimates of statistical parameters required for the sample size calculation, and (iii) to determine whether the intervention components worked together smoothly and whether there were any design issues that compromised the validity, reliability or objectivity of the planned study.

5.2 Pilot vs. feasibility

A pilot study is a study which employs, on a small scale, the same methods as a planned, larger-scale trial, in order to test whether the planned components can all work together smoothly (Porta, 2008). Feasibility work intends to assess whether planned study procedures are viable, practicable and workable, taking into account clinical, cultural, logistic, economic and ethical factors (Porta, 2008). Feasibility studies also often provide estimates of parameters needed to design the main study, e.g. for sample size calculations. The main aim of the present study was to generate data for sample size calculation. I also intended to assess practical aspects, such as whether people with undiagnosed symptoms would visit and engage with my Web application, and whether advertising strategies like Google Ad Words would help to recruit the target population. Thus this study constitutes feasibility work.

5.3 Pragmatic vs. explanatory trial

First, I needed to decide whether a pragmatic or explanatory approach was most appropriate for my aims. A pragmatic study aims to determine whether a certain intervention can improve certain variables in a given setting (Porta, 2008), whereas an explanatory study aims to determine which specific components of the intervention work and how these affect human behaviour in relation to psychological theory (Porta, 2008).

While pragmatic studies are often conducted within naturalistic settings and thus allow inferences about the effectiveness of an intervention in the ‘real World’, their findings tend to be
limited in theoretical, explanatory power (Möller, 2011; Roland & Torgerson, 1998). Explanatory studies, on the other hand, are typically conducted in experimental, laboratory-based settings which allow controlling of extraneous variables (Roland & Torgerson, 1998), but findings may be less applicable to naturalistic ‘real World’ settings (Sedgwick, 2014).

Considering the novelty of the field – Web-based, tailored interventions to trigger early help-seeking for symptoms have not, to my knowledge, been researched previously – I propose that a more explanatory approach is more suitable in this case and will contribute more significantly to scientific knowledge. I therefore decided to include several comparison groups, which allow us to specifically examine the effects of certain components of the intervention. However, by disseminating my intervention in the setting it is designated to be accessed on (the Web), and by recruiting participants who have actual, undiagnosed symptoms, rather than using fictional symptom scenarios, I suggest that my findings will be generalizable to a naturalistic ‘real World’ setting.

Thus, my primary focus in this study was explanatory, but I also included components to ensure the pragmatic aspect is taken into consideration.

5.4 Methods

5.4.1 Study design

Participants were randomised to one of four experimental groups comprising one intervention and three comparison groups (Table 5.1). Information tailoring and TPB-components are described in detail in Chapter 4.
Table 5.1 Description of the four study groups.

<table>
<thead>
<tr>
<th>Study group</th>
<th>Brief description</th>
<th>Detailed description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention (INT)</td>
<td>Tailored + TPB-components</td>
<td>Participants received the intervention described in Chapter 4: Intervention Development, i.e. tailored feedback about symptoms, risk factors, TPB-components, and a summary.</td>
</tr>
<tr>
<td>Comparison Group 1 (CG-TPB)</td>
<td>Untailed + TPB-components</td>
<td>Participants received the same information as INT participants (feedback on symptoms, risk factors, TPB-components and summary) but in a generic, untailored format.</td>
</tr>
<tr>
<td>Comparison group 2 (CG-Tall)</td>
<td>Tailored, no TPB-components</td>
<td>Participants received the same feedback as INT participants in a tailored format, excepting TPB-components, i.e. participants received tailored feedback on symptoms, risk factors, and a summary, but no TPB-components.</td>
</tr>
<tr>
<td>Usual Care (UC)</td>
<td>Typical lung cancer website</td>
<td>Participants received paraphrased information from the Roy Castle Lung Cancer Foundation’s ‘Signs and Symptoms’ webpage (the Roy Castle Lung Cancer Foundation, n.d.)</td>
</tr>
</tbody>
</table>

5.4.2 Procedure

Ethical approval was obtained from the University Research Ethics Committee (Committee 1, Ref: 15353, approval letter in Appendix 10). The study ran between October 2015 and January 2016. The study procedure is shown in Figure 5.1. On entry to the intervention homepage, users received information about the study and the Participant Information Sheet (Appendix 11). Users were informed that the website was designed not to enable self-diagnosis but to help people decide whether they should seek medical help for their symptoms, and that only those aged 18 or above and living in the UK were eligible to take part.

Following online informed consent, eligible participants completed baseline measures and were then allocated to one of four possible experimental groups using a block randomisation procedure in blocks of two. Block randomisation ensures that an equal number of participants are allocated to each study group, thus enhancing the likelihood of equally sized samples, whereas a simple randomisation procedure might assign a different number of participants to each study group. Uneven sample sizes can lead to reduced statistical power for hypothesis testing, as well as increasing the risk for bias (Efird, 2011).
Subsequently, participants received information on symptoms and risk factors which was either tailored or untailored and either included TPB components or not, depending on their study group. After clicking through the information, all participants completed the primary outcome measure (‘button click measure’, see below). If participants indicated that they would like to make an appointment to see their GP, they were, at the end of the study, directed towards the NHS online appointment booking website (NHS Choices, n.d.-d). Otherwise they were directed back to the study homepage at the end of the study. After completing the button click measure, participants had the option to end the study at this point, or to complete a further brief survey. By allowing participants to opt out of the final survey, I hoped to mitigate attrition. Attrition is a considerable issue in online studies, as the dropout threshold is very low (Webb et al., 2010).

5.4.3 Participants

I aimed to recruit both people with potential lung cancer symptoms and proxies searching on their behalf.

**Eligibility**

I recruited participants who were above the age of 18 and who were experiencing any of the following undiagnosed symptoms, or who had a friend or relative with any of the following symptoms:

- A cough
- A long-standing cough that changes or gets worse
- Dyspnoea
- Discomfort in the chest, shoulder or back
- Haemoptysis
- Hoarseness
- Unexplained weight loss or unexplained loss of appetite
- Swelling of face and/or neck
- Persistent / recurring chest infections
- Fatigue
- Finger clubbing

As information disseminated in the intervention was based on NHS guidelines and may not be applicable to other countries, I included only people who lived in the UK at time of participation (determined by self-reported geographic location).
Figure 5.1 This flowchart details the study procedure. TPB = Theory of Planned Behaviour. INT = Intervention group, CG-TPB = Comparison group 1: untailored + TPB components, CG-Tail = Comparison group 2: tailored, no TPB components, UC = usual care, based on Roy Castle website on lung cancer symptoms [6].
Recruitment

As my target sample involved people with professionally undiagnosed symptoms, I needed to adopt a broad recruitment strategy. My recruitment strategy included sending the study invitation out via mailing lists (e.g. staff and student mailing lists, mailing lists of clubs and groups such as church groups or book clubs), social media, advertising pages like gumtree and craigslist, and various websites (e.g. a website for senior citizens, and the Salford Citizen Scientist website). Furthermore, a press release was sent out via the University press office with details of the study. This resulted in various published articles as well as a radio interview.

One of my main recruitment avenues was through the Google Ad Word service. AdWords is an online advertising service which allows advertisers to have their website displayed at the top of Google search results when certain search terms (‘adwords’) are entered. This means that I was able to purchase adwords related to symptoms of lung cancer, such as “persistent cough”, “coughing up blood”, “hoarse voice”, “recurrent chest infections”, “lung cancer symptoms”, or “lung cancer warning signs”, and my study link was displayed near the top of Google search result pages when these terms were entered. My choice of Ad Words was informed by the search terms reported by participants in my survey study (Chapter 3). If a user entered “persistent cough” into Google, for example, my study link was displayed at the top of search results (denoted as an Ad).

Finally, my study was advertised by handing out leaflets to participants of a cancer awareness roadshow, and study details were also disseminated by a Tackling Cancer Together group, which aims to spread key messages about cancer to the community.

Sample size

Following recommendations by Browne (1995) for pilot studies, I aimed to recruit 30 participants per experimental group.

5.4.4 Outcome measures

I begin by presenting the measures that were completed prior to receiving treatment, and then present the measures assessed post-treatment, in the order of completion by participants.
**Initial survey**

Participants completed an initial survey prior to being randomised to one of the study groups (Appendix 12). First, participants were asked whether they were using the website for themselves, or on behalf of someone else. Depending on their response, subsequent questions were phrased accordingly. Participants were then asked whether they were experiencing any symptoms of lung cancer (or whether their friend/relative was). If a symptom was endorsed, participants were asked whether this symptom had persisted for ≥three weeks or <three weeks, and whether they experienced it as intense/severe (yes/no). Symptoms included: Cough, a change in a long-standing cough, coughing up blood, breathlessness, chest / shoulder / back pain, changes in the voice, noises when breathing, fatigue, swellings/lumps in the face or neck area, recurrent or persistent chest infections, and finger clubbing. Participants were also asked whether they had already seen a doctor about their symptoms. Participants were asked to indicate the sex, age, and smoking status (current smoker, ex-smoker or never-smoker) of the person experiencing the symptoms. This was assessed pre-treatment.

**TPB-components item**

Participants were presented with a list of statements representing the TPB constructs I targeted (Table 5.2), and asked “Which of the following statements do you agree with most?” (Appendix 13). This question aimed to identify which beliefs were most likely to impede an individual’s help-seeking behaviour. This was assessed pre-treatment.

**Table 5.2 Statements used to tailor TPB components to individual users.**

<table>
<thead>
<tr>
<th>Statement</th>
<th>TPB construct</th>
</tr>
</thead>
<tbody>
<tr>
<td>e) Making an appointment to see a doctor about these symptoms would be pointless/harmful/bad.</td>
<td>Behavioural beliefs</td>
</tr>
<tr>
<td>f) It is important whether others (family and friends) think it is necessary to see a doctor about these symptoms.</td>
<td>Normative beliefs – family/friends</td>
</tr>
<tr>
<td>g) It is important whether doctors think the symptoms are worth investigating.</td>
<td>Normative beliefs – health professional</td>
</tr>
<tr>
<td>h) Getting an appointment to report these symptoms to a doctor would be difficult.</td>
<td>Control beliefs</td>
</tr>
</tbody>
</table>
**Primary outcome (Button click behaviour)**

As I was unable to measure participants’ actual help-seeking behaviour, my primary outcome was a behavioural proxy measure: I measured whether participants clicked a button displayed at the end of the website (post-treatment), labelled “Yes” or “No” in response to the question: “Would you like to find out how to book an appointment with your doctor now?” (Appendix 14). Clicking “Yes” would redirect participants to the NHS website for making appointments online (NHS Choices, n.d.-d). This constitutes a behavioural proxy measure as it involves clicking a button to perform a certain action that is likely to lead to the target behaviour of making an appointment with a doctor. Specifically, I was interested in the proportion of patients who appropriately click this button, as help-seeking may not be appropriate in some cases. Which button click behaviour was considered appropriate or inappropriate is illustrated in Figure 5.2. This is in line with previous research (Hamilton et al., 2005) as well as guidelines developed by the UK National Institute for Health and Care Excellence for management of suspected cancer (National Institute for Health and Care Excellence, 2015).

**Intention**

Participants were asked: “Do you intend to make an appointment with your doctor to have your symptoms checked?” (Yes/No) (Appendix 15). This was assessed post-treatment.

**Demographics**

Age, gender, and ethnicity measures were adapted from the census survey (Office for National Statistics (ONS), 2011) and the measure for educational level was adapted from (Schneider, 2011) (Appendix 16). This was assessed post-treatment.
TPB questionnaire (optional)

A standardised TPB measure across different behaviours does not exist, as TPB questionnaires need to be adapted to the specific behaviour under study and the unique set of beliefs that influence the behaviour (Ajzen, 2006b). However, Ajzen (2006b) provides clear guidelines on how to construct a TPB questionnaire, and several detailed examples are available in the literature (Francis et al., 2004; Williams & Hamilton-West, 2013). Thus, the structure and wording of the questionnaire was based on recommendations by Ajzen (2006b), Williams and Hamilton-West (2013) and Francis et al. (2004), though the questionnaire was shortened to one item per construct to reduce the burden on participants following feedback from participants in the Think Aloud evaluation (Chapter 4).

Questionnaire content was based on studies on barriers to help-seeking for lung cancer symptoms as well as findings from my first study (Table 4.2 in Chapter 4). For example, Ajzen

---

*Figure 5.2 Illustration of cases in which button click behaviour is considered appropriate or inappropriate.*
(2006b) suggests that behavioural belief strength should be assessed using the following format:

*My exercising for at least 20 min, three times per week for the next three months will result in my having a faster recovery from my surgery.*

likely :___1___:___2___:___3___:___4___:___5___:___6___:___7___: unlikely

I used the same format, and used findings from previous literature and my first study to decide which beliefs to focus on. For example, previous literature shows that people with lung cancer symptoms often believe that their symptoms are not likely to be related to lung cancer. My item thus looked like this:

*Making an appointment to report these symptoms to a doctor can help to detect lung cancer early.*

likely :___1___:___2___:___3___:___4___:___5___:___6___:___7___: unlikely

In this study, not all participants may have had symptoms warranting help-seeking. To control for this, all participants were advised at the top of each page of the questionnaire: “In the following, all questions refer to what you *would* do if your symptoms persist for three weeks or longer.”

All items were measured using 7-point Likert scales. An overview of the questionnaire is provided in Table 5.3 (for the full questionnaire, see Appendix 17). This was assessed post-treatment.

*Table 5.3 Design and structure of the TPB questionnaire.*

<table>
<thead>
<tr>
<th>TPB construct</th>
<th>Number of items</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioural intention</td>
<td>1</td>
<td>“If my symptoms persisted for 3 weeks or longer, I would intend to make an appointment with my doctor to have my symptoms checked.” On a 7-point scale from ‘strongly disagree’ to ‘strongly agree’</td>
</tr>
<tr>
<td>Attitudes towards the behaviour</td>
<td>1</td>
<td>“For me to make an appointment with my doctor to have these symptoms checked would be 1 ‘pointless’ to 7 ‘useful’.”</td>
</tr>
<tr>
<td>Subcategory</td>
<td>Details</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Behavioural beliefs and their outcome evaluations</td>
<td>4 beliefs, 2 items per belief; one item to assess the perceived likelihood of the behaviour leading to a certain outcome and a corresponding item to assess the evaluation of this outcome. Beliefs assessed: “Making an appointment…” (1)… is important to check for a serious cause” (2)… can help to set your mind at rest” (3)… can help to detect lung cancer early” (4) …would waste the doctor’s time”</td>
<td></td>
</tr>
<tr>
<td>Subjective norms</td>
<td>Belief: “Making an appointment with my doctor to have these symptoms checked can help to detect lung cancer early.” from 1 ‘Very unlikely’ to 7 ‘Very likely’ Outcome evaluation: “Detecting lung cancer early is…” ‘good’ vs. ‘bad’</td>
<td></td>
</tr>
<tr>
<td>Normative beliefs and the motivation to comply with each belief</td>
<td>“Most people who are important to me want me to make an appointment with my doctor to have these symptoms checked” from 1 ‘strongly disagree’ to 7 ‘strongly agree’ Belief: “My doctors would want me to make an appointment to have these symptoms checked” from 1 ‘strongly disagree’ to 7 ‘strongly agree’ Motivation to comply: “When it comes to matters of health, I want to do what my doctors think I should do” from 1 ‘strongly disagree’ to 7 ‘strongly agree’</td>
<td></td>
</tr>
<tr>
<td>Perceived behavioural control</td>
<td>“For me to make an appointment with my doctor to have these symptoms checked would be…” 1: ‘difficult’ to 7: ‘easy’</td>
<td></td>
</tr>
</tbody>
</table>
Control beliefs and the perceived power of each belief (Note: Due to a technical error in the database, control belief items were unavailable for the feasibility study and the first trial. The error was corrected for the final trial.)

2 (1 item measured the strength of the control belief, and 1 measured the corresponding perceived power to influence the behaviour)

Belief: "I expect that my doctor's practice will not have many appointments available." from 1 'strongly disagree' to 7 'strongly agree'

Perceived power: "My doctor's practice not having many appointments available will prevent me from making an appointment to see a doctor about these symptoms." from 1 'strongly disagree' to 7 'strongly agree'

**Risk Perception scale (optional)**

Cancer risk perception was measured using a single item adapted from Rees et al. (2004): ‘How likely do you feel it is that you will develop lung cancer in your lifetime?’ rated on a 5-point Likert scale from 1, ‘very unlikely’ to 5, ‘extremely likely’ (Appendix 18). This was assessed post-treatment.

**Intervention/study group evaluation (optional)**

I measured how participants evaluated the intervention or other study group they were randomised to, by asking participants ‘Would you recommend this website to a friend if they were having similar symptoms?’ (yes/no) (Appendix 19). This was assessed post-treatment.

**Free text comments box (optional)**

Participants were invited to leave any comments they might have about the study or the website in a free text box on the final study page.

**Google Analytics**

Data on website usage were collected via Google Analytics (Google, n.d.). Google Analytics is a Web analytics service which tracks and reports website traffic for a given website. Google Analytics collects a large, complex amount of data; an overview is shown in Table 5.4.
Table 5.4 Overview of data collected by Google Analytics service.

<table>
<thead>
<tr>
<th>Category</th>
<th>Parameters measured by Google Analytics service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audience</td>
<td>Demographics (age, gender)́, interestś, geographic location based on IP address, behaviour (new user vs returning), engagement (session duration), technology (browser, network), devices (mobile, desktop, tablet, brand/model of device used)</td>
</tr>
<tr>
<td>Acquisition</td>
<td>Source/medium from which traffic to the website was directed (e.g. google, website, social media), URLs of websites from which users were referred to the website, AdWords (search terms used, no. of users referred through each keyword)</td>
</tr>
<tr>
<td>Behaviour</td>
<td>Number of page views, number of users, bounce rate, % exit (number of exits/number of page views per page), average time spent on page</td>
</tr>
</tbody>
</table>

́ estimate based on users’ previous Web interactions

5.4.5 Hypotheses

I hypothesised that the proportion of participants appropriately responding to the button-click measure would be higher among those in the intervention group than in the comparison groups.

**H1:** \( \text{INT}_{\text{app}} > \text{CG-TPB}_{\text{app}} \)

**H2:** \( \text{INT}_{\text{app}} > \text{CG-Tail}_{\text{app}} \)

**H3:** \( \text{INT}_{\text{app}} > \text{UC}_{\text{app}} \)

As the UC group received neither tailoring nor TPB components, I expected CG-TPB and CG-Tail to have a higher proportion of those clicking the button appropriately than the UC group.

**H4:** \( \text{CG-TPB}_{\text{app}} > \text{UC}_{\text{app}} \)

**H5:** \( \text{CG-Tail}_{\text{app}} > \text{UC}_{\text{app}} \)

I did not hypothesise a difference between CG-TPB and CG-Tail. The literature suggests that tailoring and addition of theory-based components might impact on behaviour, but it does not allow any conclusions whether one would be more or less effective than the other.

5.4.6 Analyses

All statistical tests were carried out using a significance level of \( \alpha = .05 \) and using the statistical software package IBM SPSS Statistics 22. The proportion of appropriate button clicks was calculated for each intervention arm, as well as the proportion who appropriately answered
yes/no on the intention item (“Do you intend to make an appointment with your doctor to have
your symptoms checked?”). Chi square tests (or Fisher’s exact test where expected frequencies
were below 5) were used to test for significant differences between study groups in the
proportion of appropriate clicks. The non-parametric Kruskal-Wallis test was used to test for
significant differences between the four groups on any continuous variables. If significant,
results of the Kruskal-Wallis test were followed up with Dunn’s Multiple Comparison post-hoc
test (Dunn, 1964).

Participants’ free-text responses in the comments box at the end of the study were analysed
using a thematic approach (Braun & Clarke, 2006). Comments were read several times to
achieve familiarisation, and to search for initial themes. Responses were then grouped
according to themes, which facilitated comparison across participants within themes and
generation of further codes.

5.5 Ethical considerations

Two main ethical considerations are of importance to this study. First, as with any health
website, it is crucial to ensure that the content is accurate, current, comprehensive and
intelligible. This is addressed in Chapter 4.

The second ethical consideration is the potential of the intervention to cause inappropriate
feelings of worry, fear or concern. I endeavoured to counteract this through the following:

- On the first page, it is clearly specified that this website is meant to help people decide
  when to seek help and should not be used to self-diagnose.
- The terms ‘cancer’ and ‘lung cancer’ are mentioned as little as possible, focusing on
  symptoms and when they warrant help-seeking rather than on possible diagnoses.
- Throughout the intervention, it is made clear that it is very unlikely that symptoms are
  caused by lung cancer, but that it is nevertheless important to have symptoms
  examined.
5.6 Results

5.6.1 Sample characteristics

Between November 2015 and January 2016, 2,463 users visited the study website. Overall, 76.0% of users left the page without any further interactions; in other words, 24% (approx. 614 users) remained on the page and undertook further interactions, such as clicking on the ‘About’ link, reading the Participant Information Sheet, or clicking the ‘consent’ button. 3,114 sessions took place (a session is a group of interactions by a user; a session ends after inactivity of 30 minutes / after midnight), with 2.96 page views per session. The majority of sessions (2,181/3,114, 70.0%) took place in the UK (based on IP address). According to Google Analytics estimates, the majority of users visiting my website were male (54.2%), and most were aged 25-34 years (Figure 5.3).

![Age distribution of visitors of the study website based on estimates by Google Analytics.](image)

In total, 130 participants completed the study. Due to a programming oversight, the database allowed us to assess only how many participants completed the study; data on how many were randomised and dropped out during the study were not available during the feasibility phase. Differential dropout occurred, with significantly more participants retained in the ‘Usual care’ group than in the other three groups, as indicated by 95% confidence intervals in Figure 5.4.
Out of 130, 116 (116/130, 89.2%) participants reported that they used the website for their own symptoms, and 14 (14/130, 10.8%) used it on behalf of someone else. According to self-reports of participants, the majority of participants were female (90/130, 69.2%), of White ethnicity (121/130, 93.1%), and education level below University level (77/130, 59.2%) (Table 5.5). The mean reported age of the sample was 49.76 years (SD =15.19), ranging from 18 to 83 years.

Table 5.5 Self-reported demographic data of the sample (N = 130).

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>40</td>
<td>30.8</td>
</tr>
<tr>
<td>Female</td>
<td>90</td>
<td>69.2</td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>3</td>
<td>2.3</td>
</tr>
<tr>
<td>Primary School</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Secondary School</td>
<td>43</td>
<td>33.1</td>
</tr>
<tr>
<td>Post Secondary School, e.g. A levels</td>
<td>31</td>
<td>23.9</td>
</tr>
<tr>
<td>Undergraduate degree</td>
<td>28</td>
<td>21.5</td>
</tr>
<tr>
<td>Postgraduate degree</td>
<td>25</td>
<td>19.2</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>121</td>
<td>93.1</td>
</tr>
<tr>
<td>Asian</td>
<td>6</td>
<td>4.6</td>
</tr>
<tr>
<td>Black</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>2</td>
<td>1.5</td>
</tr>
</tbody>
</table>
In 14 cases, the participant completing the study was not the same as the person experiencing the symptoms, as I also included people who were researching symptoms on behalf of someone else (“proxies”). While the data described above pertain to those visiting the website, I also collected information about the persons experiencing the symptoms (in the 14 cases, this information was gained by proxy). These data are shown in Table 5.6.

Table 5.6 Demographic data of those experiencing the symptoms (either self-reported or reported by proxy).

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Over 40</td>
<td>101</td>
<td>77.7</td>
</tr>
<tr>
<td>Under 40</td>
<td>29</td>
<td>22.3</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>83</td>
<td>63.8</td>
</tr>
<tr>
<td>Male</td>
<td>47</td>
<td>36.2</td>
</tr>
<tr>
<td><strong>Smoking status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never smoker</td>
<td>51</td>
<td>39.2</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>54</td>
<td>41.5</td>
</tr>
<tr>
<td>Smoker</td>
<td>25</td>
<td>19.2</td>
</tr>
</tbody>
</table>

As Figure 5.5 shows, the most commonly reported symptom was cough, followed by chest/shoulder/back pain, fatigue, dyspnoea and wheezing. More than a third of the sample (47/113, 36.2%) had reportedly already seen a doctor about the symptoms.

![Figure 5.5 Number of participants reporting various symptoms.](image-url)
5.6.2 TPB-statements

Table 5.7 shows the number and percentage of participants selecting each of the four TPB-statements. Most participants selected the TPB statement related to normative beliefs about health professionals. The statement related to control beliefs was the least frequently chosen.

Table 5.7 Number and percentage of participants selecting each of the four TPB-statements.

<table>
<thead>
<tr>
<th>TPB</th>
<th>Statement</th>
<th>n</th>
<th>%   (N=130)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioural beliefs</td>
<td>“Making an appointment to see a doctor about these symptoms would be pointless/harmful/bad.”</td>
<td>18</td>
<td>13.8</td>
</tr>
<tr>
<td>Normative beliefs – family/friends</td>
<td>“It is important whether others (family and friends) think it is necessary to see a doctor about these symptoms.”</td>
<td>23</td>
<td>17.7</td>
</tr>
<tr>
<td>Normative beliefs – health professional</td>
<td>“It is important whether doctors think the symptoms are worth investigating.”</td>
<td>81</td>
<td>62.3</td>
</tr>
<tr>
<td>Control beliefs</td>
<td>“Getting an appointment to report these symptoms to a doctor would be difficult.”</td>
<td>8</td>
<td>6.2</td>
</tr>
</tbody>
</table>

5.6.3 AdWords

Between October 2015 and January 2016 (i.e. the time frame of recruitment for this feasibility study), the study link was displayed 53,260 times via the AdWords service, and 1,721 clicks took place to the website via the Ad Words campaign (Table 5.8). The largest proportion of users who clicked on the Ad consisted of those searching for keywords that match “lung cancer” (1,016/1,721, 59.0%). Note that this includes any related search terms, such as “lung cancer symptoms” or “lung cancer diagnosis”, and any misspellings e.g. “lugn cancer”. The second largest proportion was those searching for keywords that match “Coughing up blood” (257/1,721, 14.9%), followed by those searching for “chest pains” (237/1,721, 13.8%).
Table 5.8 Number of clicks on the study link via the Google Ad Words campaign, grouped by search terms entered by users into Google.

<table>
<thead>
<tr>
<th>Term</th>
<th>Clicks</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung cancer</td>
<td>1,016</td>
<td>59.0</td>
</tr>
<tr>
<td>Coughing up blood</td>
<td>257</td>
<td>14.9</td>
</tr>
<tr>
<td>Chest pains</td>
<td>237</td>
<td>13.8</td>
</tr>
<tr>
<td>Persistent cough</td>
<td>145</td>
<td>8.4</td>
</tr>
<tr>
<td>Hoarseness</td>
<td>66</td>
<td>3.8</td>
</tr>
<tr>
<td>Recurring chest infections</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Smoking cessation</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

5.6.4 Did the study groups differ prior to receiving the intervention/controls?

A Kruskall-Wallis test showed that there was no significant difference in age between participants in the four groups ($\chi^2(3)=4.42, p=0.22$), with a mean age of 47.6 years for INT, 47.3 for CG-TPB, 54.9 for CG-Tail and 49.2 for UC. There was also no significant difference between study groups in education level ($\chi^2(3)=0.72, p=0.87$), with mean rank education level scores of 68.6 for INT, 60.9 for CG-TPB, 63.6 for CG-Tail and 67.0 for UC. Using the chi-square test, I found no significant difference between the four groups in gender, self-reported ethnicity nor smoking status (Table 5.9).

5.6.5 Button-click measure

Using the criteria shown in Figure 5.2, I created a variable denoting whether help-seeking was considered appropriate based on the participants’ responses. Help-seeking was coded as ‘appropriate’ for 98.5% (128/130) of the sample. I then created a further variable indicating whether the button was clicked appropriately, by comparing the participants’ button-click response against my variable of whether help-seeking was appropriate. If the responses matched, e.g. my variable indicated help-seeking was appropriate, and the participant clicked ‘yes’, this would be coded as appropriate, whereas a mis-match would be coded as inappropriate. Table 5.10 shows the number of appropriate button clicks per study group.
Table 5.9 Comparison between the four study groups on categorical demographic variables using chi-square test.

<table>
<thead>
<tr>
<th></th>
<th>INT count (expected count)</th>
<th>CG-TPB count (expected count)</th>
<th>CG-Tail count (expected count)</th>
<th>UC count (expected count)</th>
<th>Chi-square statistic, p-value (Asymptotic, 2 sided)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>8 (7.7)</td>
<td>6 (7.1)</td>
<td>8 (8.6)</td>
<td>18 (16.6)</td>
<td>$\chi^2(3)=0.49, p=0.92$</td>
</tr>
<tr>
<td>Female</td>
<td>17 (17.3)</td>
<td>17 (15.9)</td>
<td>20 (19.4)</td>
<td>36 (37.4)</td>
<td></td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never smoker</td>
<td>7 (9.8)</td>
<td>12 (9.0)</td>
<td>14 (11.0)</td>
<td>18 (21/2)</td>
<td>$\chi^2(6)=6.29, p=0.39$</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>11 (10.4)</td>
<td>8 (9.6)</td>
<td>11 (11.6)</td>
<td>24 (22.4)</td>
<td></td>
</tr>
<tr>
<td>Smoker</td>
<td>7 (4.8)</td>
<td>3 (4.4)</td>
<td>3 (5.4)</td>
<td>12 (10.4)</td>
<td></td>
</tr>
<tr>
<td>Self-reported ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>23 (23.3)</td>
<td>22 (21.4)</td>
<td>26 (26.1)</td>
<td>50 (50.3)</td>
<td>$\chi^2(9)=4.54, p=0.87$</td>
</tr>
<tr>
<td>Asian</td>
<td>1 (1.2)</td>
<td>1 (1.1)</td>
<td>1 (1.3)</td>
<td>3 (2.5)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0 (0.2)</td>
<td>0 (0.2)</td>
<td>0 (0.2)</td>
<td>1 (0.4)</td>
<td></td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>1 (0.4)</td>
<td>0 (0.4)</td>
<td>1 (0.4)</td>
<td>0 (0.8)</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.10 Number of people who clicked the button-click measure appropriately.

<table>
<thead>
<tr>
<th>Study group</th>
<th>n</th>
<th>Number who clicked appropriately</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT</td>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>CG-TPB</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>CG-Tail</td>
<td>28</td>
<td>4</td>
</tr>
<tr>
<td>UC</td>
<td>54</td>
<td>15</td>
</tr>
</tbody>
</table>

Using Fisher’s exact test, I found that study group allocation was not significantly related to the proportion of appropriate clicks ($\chi^2(3)=3.40, p=0.37$). Thus I did not find evidence for hypotheses H1-H5.
5.6.6 Intention

Using the approach described above, I also determined whether participants answered yes or no to the intention measure (“Do you intend to make an appointment with your doctor to have your symptoms checked?”) appropriately (Table 5.11).

Table 5.11 Number of people who responded to the intention item appropriately.

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Number who clicked appropriately</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td>CG-TPB</td>
<td>23</td>
<td>14</td>
</tr>
<tr>
<td>CG-Tail</td>
<td>28</td>
<td>15</td>
</tr>
<tr>
<td>UC</td>
<td>54</td>
<td>39</td>
</tr>
</tbody>
</table>

There was no significant relationship between study group and proportion of appropriate responses to the intention item ($\chi^2(3)=3.17$, $p=0.37$).

5.6.7 Optional measures

The optional measures were completed by 17/25 participants in the INT group, 17/23 in CG-TPB, 14/28 in CG-Tail, and 37/54 in UC. The Kruskall-Wallis test was used to test for significant differences between the four experimental groups in ‘intention to make an appointment with the doctor to have the symptoms checked’ as well as its three determinants, behavioural attitude, subjective norm, and perceived behavioural control (Table 5.12). I found that the four study groups differed significantly in scores on the intention variable, with the highest intention reported among participants in the INT group (Table 5.12). Dunn’s Multiple Comparisons test showed a significant difference between INT and CG-Tail ($p = 0.03$), i.e. between the group receiving tailored information with TPB components, and those receiving the same tailored information but without TPB components. Figure 5.6 shows a comparison of the mean intention between the four study groups.

There were no significant differences across the four study groups in any of the other TPB items (Table 5.12), and no significant differences on the cancer risk perception scale ($\chi^2(3)=1.01$, $p=0.80$).
Table 5.12 Comparison between the four study groups on TPB constructs.

<table>
<thead>
<tr>
<th>TPB construct</th>
<th>INT mean rank N=17</th>
<th>CG-TPB mean rank N=17</th>
<th>CG-Tail mean rank N=14</th>
<th>UC mean rank N=37</th>
<th>Kruskall Wallis chi-square statistic, p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention to make an appointment</td>
<td>53.6</td>
<td>42.9</td>
<td>29.4</td>
<td>43.3</td>
<td>(\chi^2(3)=8.14, \ p=0.04)</td>
</tr>
<tr>
<td>Behavioural attitude</td>
<td>46.6</td>
<td>39.3</td>
<td>33.6</td>
<td>46.6</td>
<td>(\chi^2(3)=3.87, \ p=0.28)</td>
</tr>
<tr>
<td>Subjective norm</td>
<td>41.9</td>
<td>45.9</td>
<td>43.6</td>
<td>42.0</td>
<td>(\chi^2(3)=0.47, \ p=0.93)</td>
</tr>
<tr>
<td>Perceived behavioural control: Ease of seeking help</td>
<td>35.2</td>
<td>39.9</td>
<td>44.2</td>
<td>47.5</td>
<td>(\chi^2(3)=3.39, \ p=0.34)</td>
</tr>
</tbody>
</table>

Figure 5.6 Line graph showing the mean (95% CI) intention to seek medical help (7-point scale) of the four study groups.

5.6.8 Reception and comments

After completing the final questionnaire, participants were asked whether they would recommend the website to a friend experiencing similar symptoms. The majority of those who responded to this question answered yes (75/85), and there was no significant difference between the four groups in the proportion of participants who answered ‘yes’ \(\chi^2 (3) = 4.74, \ p = 0.19\).

Participants were also invited to write a comment at the end of the final questionnaire. Thirty-one comments were made. The comments were analysed thematically and four common themes were found.
**Positive feedback**

Fourteen people commented positively on the website. The highest proportion of positive comments was in INT (5/8).

Some users reported that they had found the website useful.

> “Excellent tool, very easy to use and understandable” [INT]

Furthermore, some users reported that using the website had changed their beliefs regarding help-seeking.

> “Thank you for developing this tool; I have had a 'nagging doubt' that something might be wrong (hoping it's not lung cancer) for a while, but did not want to waste my doctor's time as I didn’t think my symptoms were 'strong enough'. Its (sic) good to get an 'independent' recommendation on whether an appointment should be made or not.” [INT]

Some users also suggested that using the website had encouraged them to mention symptoms they had not thought relevant before.

> “I was given antibiotics so feel reassured at the moment but I have noticed and complained about the changes in my fingers and nails for about two weeks but didn't mention this to the GP as I didn't think it relevant. If my cough persists after I have finished the antibiotics I will go and tell the GP about the changes to my fingers.” [INT]

**Negative feedback / Room for improvement**

Ten people suggested room for improvement. The highest proportion of negative comments was in UC (4/12).

Some users suggested that the health information provided could make people anxious.

> “I think it's a good idea to have a tool like this as it may encourage people with long standing symptoms like these to go and see their GP but caution should be used as you run the risk of making people unnecessarily anxious about their symptoms or condition.” [UC]

Several users criticised that the website did not take context (i.e. existing conditions) into account.

> “I was diagnosed with bronchiectasis [i]n 2007. Referring back 3 weeks in my life therefore does not give a true picture for the purpose of the tool which seems to be directed to people with cough and no previous diagnosis. Some of the questions therefore give an inaccurate flow and response.” [INT]

Finally, some users said the website had not answered their questions and they still felt confused about their symptoms.
“My chest feels very heavy and painful quite frightening but even though my father, brother and aunties and uncle all died of heart attacks when I was checked I was told I was fine but if it continues for a few hours might phone as my back is hurting too. This site hasn’t (sic) answered any of my question (sic) of when I should phone for help.” [UC]

**Previous experiences with health services**

Several people (n = 10) mentioned that they had already seen a GP; 5 of these were reportedly not satisfied with the advice received so far. Several people reported feeling dismissed and sent away with medication without diagnostic procedures.

“My GP is useless. When I visited him with my symptoms he was not ready to carry out further tests. I felt threatened by his rude attitude. I will never see him again even my symptoms gets worse.” [CG-TPB]

One person also reported feeling stigmatised because of smoking and that he felt his friend was receiving less medical attention for his symptoms because of his smoking habit.

“The idea is good one problem is my friend who smokes will not make an appointment as the GP bangs on about smoking, how he should pack up, what he should be doing and refers him to the cessation [sic] nurse. The symptoms [sic] of what he went with initially are overlooked while he feels like he is being preached at about his smoking habit.” [CG-TPB]

Difficulties obtaining an appointment due to limited resources were also reported.

“Tried to get urgent appointment yesterday. Fully booked for weeks. Told to go A&E!” [UC]

**Button-click measure**

Comments suggest some people did not use the button-click as they knew how to secure an appointment and did not feel the need for further assistance with this.

“I did not ask you to help me make an appointment with my doctor as it is very easy for me to do this, all I have to do is lift the telephone. Thank you for your help.” [CG-Tail]

### 5.7 Discussion

The aim of this study was to assess whether my planned study design for a randomised controlled trial of an online intervention for encouraging help-seeking for lung cancer symptoms was feasible. I aimed to test whether it would be possible to recruit the target population of people with undiagnosed symptoms potentially related to lung cancer, whether the intervention components worked together smoothly, and to obtain estimates of statistical parameters for the sample size calculation. Overall, my study was successful in recruiting the target population,
though some design issues were identified which may have compromised the validity of findings. I found some indication that components based on the TPB had causal effects on intention to seek medical advice, though the present study is underpowered to test these effects.

This study constitutes the pilot/feasibility element of the MRC Guidance for Developing and Evaluating Complex Interventions (Craig et al., 2008). This element facilitates planning of the trial needed to evaluate the intervention, and allows the identification of issues in study and intervention design. The main focus is therefore on the identification of potential issues, rather than the interpretation of changes in the outcome measures.

5.7.1 Recruiting the target population

I sought to recruit individuals who had certain symptoms but had not been diagnosed by a health professional at the time of participation. Thus I was not able to approach patients via clinical settings, support groups or other avenues usually used for patient recruitment. Instead, I had to adopt a very broad recruitment strategy, targeting the general population in order to identify people who had relevant symptoms. This involved sending out study advertisements via various channels, including mailings lists, newsletters, social media, advertising webpages, websites, news articles, leaflets, and the Google AdWord service. With these combined efforts, I was able to recruit 130 participants over a three month period.

It could be argued that a sample of 130 is small considering 2,463 unique users visited the website. The ‘bounce rate’, i.e. the number of users who leave the entry page of the website without a single interaction, was high (76%). A high bounce rate can indicate that the website was not relevant to the users (e.g. they were looking for a different type of website), that users accidentally clicked on a link, or low levels of interest or engagement for other reasons. Unfortunately, it was not possible to assess why users left the website.

Approx. 614 users remained on the website after viewing the homepage, and engaged in further interactions, indicating some interest in the website. Out of those who showed interest, 130 (21.2%) completed the study. Low participation levels between (10-12%) have been noted in previous Web-based interventions (Paul et al., 2017; Peels et al., 2012), and some even report levels as low as 0.24% (Koo & Skinner, 2005), though others report more positive experiences (Murray et al., 2009). Additionally, online trials have often been associated with low retention
levels (Mathieu, McGeechan, Barratt, & Herbert, 2013; Murray et al., 2009). A meta-analysis of tailored, Web-based health interventions suggests, however, that high attrition rates do not necessarily bias study outcomes (Lustria et al., 2013).

Google Analytics data of the study website suggest that paid Google advertisements led to a large number of page visits, particularly from those searching for keywords like ‘lung cancer symptoms’ and ‘coughing up blood’. This shows how online platforms make it possible to reach broad audiences; by using the AdWord campaign, my study link was displayed 53,260 times and clicked on over 1000 times. Using more traditional recruitment strategies, it would be extremely difficult to reach such a large audience.

In order to assess whether the recruited participants are similar to patients with lung cancer, it is important to assess symptomatology as well as demographic risk factors like age and smoking status.

Among lung cancer patients, the most commonly reported symptoms upon presentation are cough, systematic symptoms like fatigue and weight loss, dyspnoea, chest pain, haemoptysis, symptoms of local or distant dissemination like hoarseness and stridor/wheezing, and recurring chest infections (Buccheri & Ferrigno, 2004). As Figure 5.5 shows, these symptoms were all represented in my sample, suggesting that I succeeded in recruiting individuals with relevant symptoms.

Furthermore, the average age of the sample was $M = 49.76$, and 77.7% of the sample were either aged above 40 years, or searching on behalf of someone aged above 40 years. Again, this suggests I recruited the relevant target group, as NICE guidelines for suspected lung cancer referral state that individuals with certain symptoms who are above the age of 40 years should be referred for urgent chest X-rays (National Institute for Health and Care Excellence, 2015).

The majority of participants were reportedly current or former smokers (60.7%), again indicating that the website attracted and retained those potentially at risk for lung cancer, although a higher proportion of smokers would be expected among a sample of people with lung cancer.

Moreover, it should be noted that based on the criteria shown in Figure 5.2, help-seeking was considered appropriate in 98.5% (128/130) of the sample. This suggests that my study attracted and engaged individuals with prolonged or severe symptoms that require further attention.
rather than individuals with mild, short-term symptoms. This may be due to recruitment via Google AdWords, as research suggests those using the Web to research symptoms tend to research long-term symptoms and conditions while those with acute symptoms are more likely to use telephone triage (North et al., 2011).

5.7.2 Changes for the full trial

The feasibility study highlighted some issues with the design and content of the trial, and the website, which needed to be addressed prior to launching the full online randomised controlled trial.

Differential dropout

The main design issue identified in this feasibility study pertains to the differential dropout of participants across the four study groups. Differential dropout can limit the validity of findings (Moher et al., 2010), particularly when attrition occurs for systematic rather than random reasons (Bell, Kenward, Fairclough, & Horton, 2013), e.g. when study arms differ in acceptability, as is suspected in my study.

The UC group retained approximately twice as many participants as the remaining three groups, and made up 41.5% of the final sample. I hypothesise that differential dropout occurred due to differing lengths of the four conditions. While INT, CG-TPB and CG-Tail contained several information pages – symptom information, risk factors information, TPB-components, and the summary page – the UC group involved only one page with considerably less detail. Research has shown that Internet users prefer brief, concise information for Web-based health interventions (Yardley, Morrison, Andreou, Joseph, & Little, 2010).

I decided to terminate data collection when each of the three remaining groups had ≈25 participants rather than continue until 30 was reached in each group. A continuation of data collection would have further inflated the size of the UC group, while recruitment for the other three groups would have been relatively slow, and the data would have been of compromised validity.

To mitigate the risk of differential attrition in the full trial, I needed to ensure all study groups received approximately the same number and length of information pages. I therefore decided to reduce all study groups to two information pages for the full trial – one page containing
symptom information, risk factor information and TPB components, and one page providing the summary. I also needed to ensure the four groups were similar in the amount of information provided and therefore decided to replace the UC group with a group more similar to the remaining three groups – this is described in the following section.

Another issue regarding attrition was that a programming oversight meant that the database only recorded how many participants completed the study, but did not show how many were randomised. Thus, the code was revised to create a database for the full trial which also recorded how many participants were randomised.

**Factorial design**

Following analyses of the feasibility data, it was decided that a factorial design would provide important information that could not be gained from the current design. My data showed that participants in the tailored, TPB-based group (the intervention) reported significantly higher intention to seek medical help than participants in the tailored group without TPB-components (CG-Tail). The information provided to participants in these two groups was the same information, tailored in the same way; the only difference between these two groups was that INT included TPB-components, which CG-Tail did not. This suggests TPB-components had a causal effect on intention to seek help.

However, participants who received TPB components but no information tailoring (CG-TPB) did not report significantly higher intention to seek help than participants who received tailored information without TPB-components (CG-Tail). This suggests not only the presence of TPB-components, but also information tailoring may have played a role. Whether there is an interaction between tailoring and TPB-components can be assessed using a factorial 2x2 design, with tailoring and presence of TPB-components as the two factors.

For the full trial, the UC condition was therefore replaced by a condition that involved the same type of information (symptoms, risk factors, summary) and the same number of pages (2), but in an untailored version without TPB components (CG-none). This change had two consequences: (i) this created a 2x2 design with the two factors ‘tailoring’ and ‘TPB-components’, and (ii) the four study groups were more equal in amount of information and number of pages, with the aim of mitigating differential dropout.
The UC group was replaced rather than retained for pragmatic reasons; recruitment across five groups would have added a considerable amount of time and cost in terms of recruitment, which was not deemed feasible for the present project.

**Primary outcome measure**

Only 13 people (0.1%) across all four groups clicked ‘Yes’ on the ‘button-click’ measure, although 65.4% (85/130) indicated intention to make an appointment with their doctor to have their symptoms checked. This indicates that the button-click measure might not be an appropriate measure to assess help-seeking behaviour. Moreover, participants’ comments in the comments section suggest that some people did not click this button because they did not feel they needed any additional support in making an appointment. Again, this shows that this measure has limited utility.

Therefore, I decided to change the primary outcome measure to ‘intention to make an appointment’. Many studies have provided evidence for the value of predicting behaviour using intentions, though results vary depending on the behaviour under study: For some behaviours such as condom-use and exercise, meta-analyses show correlations between intentions and behaviour that range from 0.44 to 0.56 (Ajzen, Albarracin, & Hornik, 2007; Albarracin, Johnson, Fishbein, & Muellerleile, 2001; Godin & Kok, 1996; Hausenblas, Carron, & Mack, 1997; Sheeran & Orbell, 1998), and individual studies have shown high correlation above 0.80 for behaviours like using birth control pills or breast vs. bottle feeding (Ajzen & Fishbein, 1980).

Thus, while intention does not explain the full variance in behaviour, it has been found to be a valid proxy measure.

As I still wished to assess which actions participants would undertake following completion of the study, I retained this outcome measure, but amended it in order to accommodate other options.

**Supporting people who have already sought help**

Participants’ comments showed that some users of the website had already sought medical advice, and were dissatisfied with the care they received. Users described feeling dismissed by their GPs and being sent away without further investigation of symptoms. In order to support these individuals in determining whether their symptoms required further investigation, I decided to highlight the NICE clinical guidelines provided in the summary. The guidelines state under
which conditions patients should be placed on a suspected cancer referral pathway and referred for an urgent chest X-ray. Thus this information may provide useful information to patients that they can discuss with their GP in order to decide whether they should be referred for further investigation. While this information was provided in the original intervention, it was only briefly mentioned on the summary page and was easy to overlook. For the full trial, I therefore decided to highlight this information through visual design, i.e. larger and bold font size, coloured background (Figure 5.7).

![Screen shot showing the information on NICE clinical guidelines with highlighted text.]

Figure 5.7 The information on NICE clinical guidelines was visually highlighted for the full trial, to ensure this information would be read by users, as this information could be useful to individuals who have already sought help.

**Co-morbidities**

As several users pointed out that existing conditions should be taken into account, I decided to include a measure of respiratory co-morbidities, and include any endorsed co-morbidities on the summary page. I did not change the advice provided on the Website based on co-morbidities mentioned, however, because research has shown that symptoms of lung cancer are often masked by existing conditions, making it difficult for the patient to distinguish between existing and new symptoms (Corner & Brindle, 2010; Tod & Joanne, 2010). I would therefore not want to discourage individuals with existing conditions with related symptoms from seeking medical advice.

5.7.3 Preliminary findings

I conducted some preliminary analyses to test the effects of tailoring and TPB-components. I briefly discuss these results, though the focus of the study was on feasibility aspects and sample size calculation and one should bear in mind that the present study was underpowered to draw conclusions about the wider population.
While I found no significant differences between study groups on my primary outcome measure, this may be because the measure did not accurately capture participants’ intentions or behaviours, as discussed above.

I did, however, find a significant difference between INT and CG-Tail on intention to make an appointment. This finding is interesting because INT and CG-Tail were identical excepting only the TPB components. Thus this finding suggests a causal role of TPB-components in intention to seek medical advice. The TPB-components consisted of quotes, either by health professionals or (fictional) family members, and a list of specific advice on how to seek medical advice. Participants received one of these components, depending on how they responded to an initial item designed to elicit which TPB beliefs were most likely to influence their decision to seek medical advice. In INT, 16/25 participants received quotes from health professionals to target behavioural attitude, 4/25 received quotes from health professionals to target subjective norms, 4/25 received the list of steps designed to target perceived behavioural control, and one participant received quotes from fictional family members. Thus, the majority of participants in INT (20/25) received quotes from health professionals, which suggests these impacted most strongly on reported intention to seek medical advice.

I found no significant differences between the untailored study group (CG-TPB) and the tailored groups (INT and CG-Tail) on any of the included outcome measures.

5.7.4 Sample size calculation

As I planned to conduct a study with a 2x2 factorial design, I based my sample size calculation on a 2-way ANOVA with two factors with two levels each, using nQuery, based on the formulas in (O’Brien & Muller, 1993). As I decided to use intention as my primary outcome measure for the full trial, I used the means derived from the intention item from the TPB questionnaire. As my feasibility study did not include an untailored, non-TPB based group, I used the mean for the UC group as a proxy, as this condition was untailored and did not include the TPB-components (Table 5.13). The pooled standard deviation across all four groups was 2.179.
Table 5.13 Mean intention (7-point scale) by study group, organised in a 2x2 matrix with Factor A = ‘Presence of TPB components’, and Factor B = ‘Tailoring’.

<table>
<thead>
<tr>
<th></th>
<th>Tailored</th>
<th>Untailored</th>
<th>Marginal means B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TPB</strong></td>
<td>5.75 [INT]</td>
<td>4.68 [CG-TPB]</td>
<td>5.215</td>
</tr>
<tr>
<td><strong>No TPB</strong></td>
<td>3.56 [CG-Tail]</td>
<td>5.11 [UC]</td>
<td>4.335</td>
</tr>
<tr>
<td><strong>Marginal means A</strong></td>
<td>4.655</td>
<td>4.895</td>
<td>4.775</td>
</tr>
</tbody>
</table>

These data were entered into nQuery with a test significance level of 0.05. As Table 5.14 shows, a sample size of $n = 49$ per group ($N = 196$) would ensure a minimum of 80% power for a main effect for Factor A (TPB/no TPB) and for the interaction term, which is considered sufficient to detect existing effects (Cohen, 1988; Wilson Van Voorhis & Morgan, 2007).

Statistical power for Factor B (tailoring) was low as there was very low variance in means for this factor; to increase the power to at least 80% would require a large sample size that was beyond the scope of the project.

Table 5.14 NQuery output, based on feasibility data, showing the power obtained for main and interaction effects for three different possible sample sizes.

<table>
<thead>
<tr>
<th></th>
<th>Option 1</th>
<th>Option 2</th>
<th>Option 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test significance level</strong></td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>Number of Factor A levels</strong></td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Number of Factor B levels</strong></td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Variance in means, $V_A$</strong></td>
<td>0.194</td>
<td>0.194</td>
<td>0.194</td>
</tr>
<tr>
<td><strong>Variance in means, $V_B$</strong></td>
<td>0.014</td>
<td>0.014</td>
<td>0.014</td>
</tr>
<tr>
<td><strong>Variance in means, $V_{AB}$</strong></td>
<td>0.429</td>
<td>0.429</td>
<td>0.429</td>
</tr>
<tr>
<td><strong>Common standard deviation, $\sigma$</strong></td>
<td>2.179</td>
<td>2.179</td>
<td>2.179</td>
</tr>
<tr>
<td><strong>Power for A (%)</strong></td>
<td>59</td>
<td>48</td>
<td>80</td>
</tr>
<tr>
<td><strong>Power for B (%)</strong></td>
<td>9</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td><strong>Power for AB (%)</strong></td>
<td>90</td>
<td>80</td>
<td>98</td>
</tr>
<tr>
<td><strong>n per group</strong></td>
<td>30</td>
<td>23</td>
<td>49</td>
</tr>
</tbody>
</table>

Factor A: TPB / no TPB components
Factor B: Tailoring / no tailoring
5.8 Summary and conclusions

From this study, I can conclude that it is possible to recruit participants with undiagnosed symptoms for a help-seeking intervention using online recruitment strategies, and that paid advertisements such as Google Ad Words are useful in this endeavour. This information will be of use to future help-seeking and early detection trials. It will also be relevant to studies examining Web use for symptom appraisal; my systematic review suggests that experimental studies tend to use fictional symptom scenarios (Mueller, Jay, Harper, Davies, et al., 2017), but my study shows that it is possible to investigate Web use prior to diagnosis in a more ecologically valid setting.

Furthermore, this study provides crucial insights into design issues that could compromise the validity of the findings, which helped to inform amendments for the subsequent online randomised trial. It also informed my sample size calculation, ensuring relevant estimates of statistical parameters were available to perform an informed sample size calculation.

In the next chapter, I present the methods and results of the revised online trial, taking into account findings of the feasibility study.
Chapter 6: Trial 1: A randomised controlled trial to test a tailored, theory-based intervention to encourage help-seeking for lung cancer symptoms

6.1 Aim

After implementing the changes to the study design following the feasibility work, and after completing the power analysis to calculate the required sample size (Chapter 5), I conducted the full randomised trial. The aim of this study was to assess whether the TPB-based components and the information tailoring employed in the intervention had an effect on individuals’ intention to seek help for symptoms related to lung cancer. I aimed to test for main effects of both these factors as well as their interaction. Note, however, that the statistical power to detect effects of tailoring was low (see ‘Sample size calculation’ in Chapter 5). This chapter addresses the evaluation element of the MRC Guidance (Craig et al., 2008).

6.2 Methods

6.2.1 Study design

This study employed a 2x2 design, with the two factors ‘presence of TPB-components’ (yes/no) and ‘information tailoring’ (yes/no) (Table 6.1). Information tailoring and TPB-components are described in detail in Chapter 4.

Table 6.1 Study groups.

<table>
<thead>
<tr>
<th>TPB-components: Yes</th>
<th>Information tailored: Yes</th>
<th>Information tailored: No</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT</td>
<td>CG-TPB</td>
<td></td>
</tr>
<tr>
<td>CG-Tail</td>
<td>CG-none</td>
<td></td>
</tr>
</tbody>
</table>

6.2.2 Participants

Eligibility and recruitment

Eligibility criteria as well as recruitment strategies were the same as those used in the feasibility study (Chapter 5).
Sample size

Based on the power analysis reported in Chapter 5, I aimed to recruit a minimum of 49 participants per study group who completed the website and provided outcome measurement.

6.2.3 Outcome measures

Initial survey

The initial survey was expanded to include an item asking participants whether they had any respiratory co-morbidities (Appendix 20). Questions on symptoms and risk factors were the same as in the feasibility study (Chapter 5).

TPB-components item

This item remained the same as in the feasibility study (Appendix 13).

Intention (primary outcome)

The primary outcome for the trial was reported intention to seek medical help (as a proxy measure of participants’ help-seeking behaviour). Participants responded to the question: “Do you intend to see a doctor about your symptoms?” on a 7-point scale from 1: “No, definitely not” to 7: “Yes, definitely.” Proxies were asked: “Do you intend to encourage your friend/relative to see a doctor about their symptoms?” (Appendix 21).

‘Next steps’ measure

The ‘button-click’ measure was expanded to include other options and thus renamed ‘Next steps’ measure (Appendix 22). After viewing the information pages, participants were asked: “What would you like to do now?” and presented with four clickable buttons, labelled respectively:

1. Make an appointment with my GP
2. Register with a GP
3. Change to a different GP
4. None of these

Depending on the button clicked, participants were redirected to relevant NHS information pages after study completion.
Demographics

The demographics measure was the same as in the feasibility study (Chapter 5).

Optional survey

The TPB questionnaire, the Risk Perception scale and the Intervention/study group evaluation item were the same as in the feasibility study (Chapter 5).

Google Analytics

Data on website usage were collected via Google Analytics (details in Chapter 5).

6.2.4 Procedure

Ethical approval was obtained from the University Research Ethics Committee (Committee 1, Ref: 15353, approval letter in Appendix 10). The study ran between February 2016 and June 2016. The study procedure is shown in Figure 6.1. The procedure was largely the same as for the feasibility study (Chapter 5), though one of the study groups (‘Usual Care’) was exchanged for “CG-none”, an untailored version without TPB-components. Furthermore, participants in the trial received symptom and risk factor information as well as TPB-components on the same page, whereas these were displayed on separate pages in the feasibility study. Finally, some changes were introduced to the post-treatment battery of measures, as the ‘button-click’ measure was changed into the ‘next-steps’ measure, and the primary outcome was now intention to seek help. The Participant Information Sheet is shown in Appendix 23.
Figure 6.1 Study procedure. TPB = Theory of Planned Behaviour. INT = Intervention group. CG-TPB = Comparison group 1: untailored + TPB-components. CG-Tail = Comparison group 2: tailored, no TPB-components. CG-none = Comparison group 3: untailored, no TPB-components. Changes from the feasibility study to the trial are highlighted in green font.
6.2.5 Data handling and coding

I wanted to take into account whether participants responded to the intention variable appropriately, as the intervention aimed to increase help-seeking intention particularly in those where help-seeking is appropriate. I therefore created a binary ‘appropriateness’ variable.

I considered responses to be appropriate if they signalled high intention to seek help and help-seeking was classed as warranted, or if they signalled low intention to seek help, and help-seeking was classed as not warranted (Table 6.2). Intention was considered high if scores were above the midpoint of the 7-point scale (>4), and as low if scores were equal to or below the midpoint (≤4). The rationale for this cut-off is that the midpoint of the scale suggests a neutral stance towards help-seeking. Thus participants scoring ≤4 can be considered either undecided towards help-seeking, or tending towards not seeking help. Scores above 4, on the other hand, suggest a tendency towards help-seeking.

Table 6.2 The table shows when responses to the primary intention variable were considered as appropriate or inappropriate.

<table>
<thead>
<tr>
<th>Help-seeking warranted:</th>
<th>“High intention to seek help”: intention &gt; 4</th>
<th>“Low intention to seek help”: intention ≤ 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Symptoms persist for ≥ 3 weeks OR</td>
<td>Appropriate</td>
<td>Not appropriate</td>
</tr>
<tr>
<td>• Symptoms are intense/severe OR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Haemoptysis is present OR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Finger clubbing is present OR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Persistent chest infections present</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Help-seeking not warranted:</th>
<th>“High intention to seek help”: intention &gt; 4</th>
<th>“Low intention to seek help”: intention ≤ 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Symptoms persist for &lt; 3 weeks AND</td>
<td>Not appropriate</td>
<td>Appropriate</td>
</tr>
<tr>
<td>• Symptoms are not intense/severe AND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Haemoptysis is not present AND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Finger clubbing is not present AND</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Persistent chest infections not present</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
For the TPB questionnaire, I created weighted belief measures, by multiplying each belief variable by the score representing the influence of the belief (Francis et al., 2004). Behavioural belief variables were multiplied by their outcome evaluation, and normative beliefs were multiplied by their motivation to comply to give a score for subjective norm (i.e. perceived social pressure to perform the behaviour).

To illustrate, a participant scoring ‘6’ on the 7-point scale for the behavioural belief item “Making an appointment to see a doctor about these symptoms, even if they are unlikely to be caused by lung cancer, can help put your mind at rest.” (1: Unlikely to 7: Likely)” and a ‘3’ on the 7-point scale for the corresponding outcome evaluation item “Putting your mind at rest is… (1: bad to 7: good)”, would receive a composite score of 6x3=18.

Control beliefs would be multiplied by perceived power, but as indicated in Chapter 5, a technical error in the database resulted in loss of data for these items in the feasibility study and this trial. I did, however, collect data on two PBC items. PBC is a product of control beliefs and perceived power (Francis et al., 2004). To create a composite PBC measure, I computed the average of these two items.

6.2.6 Hypotheses

**H1:** I hypothesised a main effect of TPB-components, such that participants who received TPB-components would report higher intention to see a doctor about their symptoms than participants who did not receive TPB-components.

**H2:** I hypothesised a main effect of tailoring, such that participants who received tailored information would report higher intention to see a doctor about their symptoms than participants who received untailored information.

**H3:** I hypothesised an interaction effect between TPB-components and tailoring, such that a combination of the two factors (INT) would result in higher intention compared to each factor alone (CG-TPB and CG-Tail), or none of the factors (CG-none).

**H4:** I hypothesised that the proportion of participants responding appropriately to the intention item would be larger in the groups receiving tailored information, compared to untailored information, because the tailored study groups provided users with individualised information on
whether they should seek medical advice, whereas the untailored groups provided generic information.

6.2.7 Statistical analyses

All statistical tests were carried out using a significance level of $\alpha = .05$ and using the statistical software package IBM SPSS Statistics 22 and R version 3.4.1. Because standardised residuals were non-normally distributed, I conducted non-parametric two-way ANOVAs using the Sheirer-Ray-Hare test (Holmes, Moody, Dine, & Trueman, 2016) to test for effects of the two factors on the primary outcome measure, the variables measured in the TPB questionnaire and lung cancer risk perception. The Sheirer-Ray-Hare test requires equal sample sizes across study groups. I therefore randomly sampled the largest possible, equal group size from each group to run the test. I repeated random sampling five times to test the robustness of results, as this is the recommended number of imputed data sets when conducting multiple imputations for missing data (Fichman & Cummings, 2003) (though in this case data were not missing but intentionally removed to create equal group sizes).

The Kruskall-Wallis test was used to test for differences in means across the study groups. I used the non-parametric Mann-Whitney U-test to test for differences between two means.

In my analyses, I also wanted to take into account whether participants responded to the intention variable appropriately. I therefore compared the proportion of appropriate/inappropriate responses to the intention item across the two factors using Pearson’s chi-squared tests.

I was also interested in identifying other variables that are related to intention to seek help in lung cancer, as this has not been assessed previously. I therefore computed Pearson correlations between intention and demographic variables as well as symptoms. For exploratory purposes, variables that were related to intention in bivariate comparisons were entered into a linear regression model to predict intention. Because the residuals of the intention variable were non-normally distributed, I used a robust regression analysis based on MM-type estimators (Maechler et al., 2016).

To test whether the study groups differed on demographic variables (i.e., to test whether the groups differed significantly prior to treatment), I tested for associations between study groups and demographic variables using the chi-squared test for categorical variables and the Kruskal-Wallis test for non-normally distributed or ordinal-scaled variables.
6.3 Results

6.3.1 Sample characteristics

Between February 2016 and July 2016, 5,004 users visited the study website, and 82.2% of users left the page without any further interactions; i.e. 17.8% (approx. 891 users) remained on the page and undertook further interactions (e.g. reading the Participant Information Sheet, or clicking the ‘consent’ button). 5,534 sessions took place (a session is a group of interactions by a user; a session ends after inactivity of 30 minutes / after midnight), with 1.93 page views per session. The majority of sessions (4,137/5,004, 82.7%) took place in the UK (based on IP address).

It should be noted that an unusually large number of sessions from Russia was documented (1,035), which may mean that the website was accessed by ‘Web crawlers’ or ‘Web robots’, which are automated programs that browse the Web. Thus, the Google analytics data may not reflect only human use of the Website. Incorporating a CAPTCHA tool would have prevented robots from proceeding through the site, as this requires users to input some information (e.g. letters from a distorted image), which would not be discernible by a robot. However, this would have required a CAPTCHA tool on every individual page of the website, greatly reducing its usability. In order to strike a balance between usability and security it was therefore decided to incorporate other mechanisms to reduce robot interactions on the site. These mechanisms meant that robots could still visit the website (and affect google analytics data), but could not feed data into the database, thus retaining the validity of the data collected via the database.

In total, 707 participants were randomised, and 212 (30.0%) participants completed the study (Table 6.3).

Table 6.3 Number and percentage of participants who were randomised and completed the study per group.

<table>
<thead>
<tr>
<th>Study group</th>
<th>n randomised</th>
<th>n completed</th>
<th>% completed</th>
<th>% of total N = 212</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT</td>
<td>178</td>
<td>53</td>
<td>29.8</td>
<td>25</td>
</tr>
<tr>
<td>CG-TPB</td>
<td>177</td>
<td>50</td>
<td>28.2</td>
<td>23.6</td>
</tr>
<tr>
<td>CG-Tail</td>
<td>176</td>
<td>50</td>
<td>28.4</td>
<td>23.6</td>
</tr>
<tr>
<td>CG-none</td>
<td>176</td>
<td>59</td>
<td>33.5</td>
<td>27.8</td>
</tr>
</tbody>
</table>
Out of 212, 189 (189/112, 89.2%) participants used the website for their own symptoms ("affected individuals"), and 23 (23/212, 10.8%) used it on behalf of someone else ("proxies"). Demographic data are shown in Table 6.4. The majority of participants were, according to self-reports, female (133/212, 62.7%), of White ethnicity (203/212, 95.8%), and education level below University level (146/212, 68.9%), and reported no co-morbidities (147/212, 69.3%). Asthma (29/212, 13.7%) and COPD (16/212, 7.5%) were the most commonly reported co-morbidities (Table 6.4).

Figure 6.2 shows the frequency of reported symptoms and the reported duration for each symptom, where duration was assessed (for symptoms like haemoptysis, where help-seeking is recommended regardless of duration, this was not assessed). More than a third (78/212, 36.8%) had already seen a doctor about their symptoms prior to study participation.

![Figure 6.2 Number of participants reporting each symptom, and proportion reporting a duration of ≥3 weeks, or less.](image-url)
Table 6.4 Demographic data of the sample (N = 212).

<table>
<thead>
<tr>
<th>Demographic Category</th>
<th>n</th>
<th>% (out of 212)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>M = 53.8, SD = 14.3, 18 – 89 years</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>79</td>
<td>37.3%</td>
</tr>
<tr>
<td>Female</td>
<td>133</td>
<td>62.7%</td>
</tr>
<tr>
<td>Education level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>2</td>
<td>0.9</td>
</tr>
<tr>
<td>Primary School</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Secondary School</td>
<td>38</td>
<td>17.9</td>
</tr>
<tr>
<td>Post Secondary School, e.g. GCSE</td>
<td>44</td>
<td>20.8</td>
</tr>
<tr>
<td>Further education, e.g. A-levels</td>
<td>61</td>
<td>28.8</td>
</tr>
<tr>
<td>Undergraduate degree</td>
<td>43</td>
<td>20.3</td>
</tr>
<tr>
<td>Postgraduate degree</td>
<td>23</td>
<td>10.8</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>203</td>
<td>95.8</td>
</tr>
<tr>
<td>Asian</td>
<td>4</td>
<td>1.9</td>
</tr>
<tr>
<td>Black</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Mixed</td>
<td>2</td>
<td>0.9</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>2</td>
<td>0.9</td>
</tr>
<tr>
<td>Co-morbidities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>147</td>
<td>69.3</td>
</tr>
<tr>
<td>Pulmonary fibrosis</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Asthma</td>
<td>29</td>
<td>13.7</td>
</tr>
<tr>
<td>COPD</td>
<td>16</td>
<td>7.5</td>
</tr>
<tr>
<td>Respiratory infection</td>
<td>7</td>
<td>3.3</td>
</tr>
<tr>
<td>Sleep apnoea</td>
<td>7</td>
<td>3.3</td>
</tr>
<tr>
<td>Cystic fibrosis</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
<td>4.2</td>
</tr>
<tr>
<td>Missing/don’t know</td>
<td>6</td>
<td>2.8</td>
</tr>
</tbody>
</table>

6.3.2 TPB-statements

As in the feasibility study, the TPB statement related to normative beliefs about health professionals was the most frequently selected option (Table 6.5). A larger proportion chose the statement related to control beliefs (16.5%) than in the feasibility study (6.2%).
Table 6.5 Number and percentage of participants selecting each of the four TPB-statements (N=212).

<table>
<thead>
<tr>
<th>TPB statement</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioural attitude</td>
<td>42</td>
<td>19.8</td>
</tr>
<tr>
<td>Social norm – Family/friends</td>
<td>34</td>
<td>16.0</td>
</tr>
<tr>
<td>Social norm – Doctors</td>
<td>101</td>
<td>47.6</td>
</tr>
<tr>
<td>Perceived behavioural control</td>
<td>35</td>
<td>16.5</td>
</tr>
</tbody>
</table>

6.3.3 Primary outcome

Overall, 212 participants completed the intention measure (M=5.2, SD=2.03).

Intention and TPB-components / tailoring

Figure 6.3 shows the means for the primary outcome measure, intention to see a doctor about the symptoms, across the four study groups. The overlapping 95% confidence intervals suggest no significant differences between the study groups in mean intention, which was confirmed by the Kruskal-Wallis test ($\chi^2(3)=0.65$, $p=0.89$). Figure 6.4 and Figure 6.5 show the mean intention for those who received TPB-components vs. no TPB-components, and tailored vs. untailored information, respectively, again with overlapping confidence intervals indicating no significant differences.

Figure 6.3 Mean (95% CI) intention to see a GP about the symptoms (measured on 7-point scale) across the four study groups.
The standardised residuals for intention were non-normally distributed (Kolmogorov-Smirnov test: $D(212)=0.25, p<0.001$) with a strong positive skew (Figure 6.6), thus violating assumptions underlying parametric ANOVA. I therefore conducted the non-parametric Sheirer-Ray-Hare test. As the Sheirer-Ray-Hare test requires equal sample sizes across the groups, I randomly selected 50 participants (the largest possible number across the four groups, Table 6.3) from each group. To achieve a replicable random sample, I set the seed to the current date, e.g. 20012017 for 20. January 2017. Using the randomly selected sample of $N=200$, I then conducted the Sheirer-Ray-Hare test in Excel2013 (Holmes et al., 2016).
Figure 6.6 Distribution of standardised residuals for the primary outcome variable, intention to see a GP about the symptoms.

As Table 6.6 shows, the Sheirer-Ray-Hare test revealed no significant main or interaction effects (neither in this random sample nor in the four other randomly selected samples, see Appendix 24). This indicates participants’ reported intention to see a GP about their symptoms was not affected by the presence of TPB-components or tailoring. Thus, the results do not support hypotheses H1-H3.

Table 6.6 Results for the Sheirer-Ray-Hare test using a randomly selected sample of n=50 per group (N=200).

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Sum of squares</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPB/no TPB</td>
<td>40.50</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>0.908</td>
<td></td>
</tr>
<tr>
<td>Tailored/Untailored</td>
<td>2380.50</td>
<td>1</td>
<td>0.79</td>
<td>3.84</td>
<td>0.373</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>1.28</td>
<td>1</td>
<td>0.00</td>
<td>3.84</td>
<td>0.984</td>
<td></td>
</tr>
</tbody>
</table>

Using the chi-squared test, I tested for associations between the binary variable denoting whether the intention variable was answered appropriately (yes/no) and tailoring; no association was detected ($\chi^2(1)=2.29, p=0.13$). Thus I did not find evidence for hypothesis 4.

**Intention and demographic variables**

As Table 6.7 shows, age was significantly associated with intention to seek help, with reported intention increasing with higher age. Smoking status and education level were not associated with intention, but there is a slight trend towards significance for smoking status.
Table 6.7 Correlations between demographic variables and primary outcome intention.

<table>
<thead>
<tr>
<th></th>
<th>Pearson Correlation</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking status (1: never / 2: ex / 3: current smoker)</td>
<td>0.12</td>
<td>0.08</td>
</tr>
<tr>
<td>Education</td>
<td>-0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>Age</td>
<td>0.25</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Note: Significant correlations are highlighted in bold.

**Intention and symptoms**

As Table 6.8 shows, intention significantly increased with presence of cough, a change in an existing cough, breathlessness, wheezing, recurring chest infection, and if the participant had already seen a doctor. All correlations were small according to Cohen’s conventions (Cohen, 1988). Interestingly, intention significantly decreased if weight loss was present, but the correlation was very small.

Table 6.8 Correlation between intention to seek help and presence of symptoms.

<table>
<thead>
<tr>
<th></th>
<th>Pearson Correlation</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cough</td>
<td>0.13</td>
<td>0.05</td>
</tr>
<tr>
<td>Change in an existing cough</td>
<td>0.21</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Haemoptysis</td>
<td>-0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>Breathless</td>
<td>0.29</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Discomfort in chest/shoulder/back</td>
<td>0.15</td>
<td>0.03</td>
</tr>
<tr>
<td>Hoarseness</td>
<td>0.04</td>
<td>0.54</td>
</tr>
<tr>
<td>Wheezing</td>
<td>0.26</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Weight loss</td>
<td>-0.14</td>
<td>0.04</td>
</tr>
<tr>
<td>Fatigue</td>
<td>0.07</td>
<td>0.30</td>
</tr>
<tr>
<td>Swelling</td>
<td>0.03</td>
<td>0.71</td>
</tr>
<tr>
<td>Recurring chest infections</td>
<td>0.14</td>
<td>0.04</td>
</tr>
<tr>
<td>Finger clubbing</td>
<td>0.07</td>
<td>0.35</td>
</tr>
<tr>
<td>Already seen a doctor?</td>
<td>0.25</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Number of respiratory comorbidities</td>
<td>0.20</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

Note: Significant correlations are highlighted in bold.
Regression model

All demographic and symptom variables significantly correlated with intention were entered into the robust regression model, resulting in the model shown in Table 6.9, which accounted for 20% of variance in intention (adjusted $R^2=0.20$). As Table 6.9 shows, age was the only significant predictor of intention to seek help, though a trend towards significance may be indicated for the presence of discomfort in the chest/shoulder/back, and for whether participants had already seen a doctor about their symptoms.

Table 6.9 Regression model of demographic and symptom variables predicting intention to seek help.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unstandardised coefficient estimate</th>
<th>Standard error</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.04</td>
<td>0.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cough</td>
<td>0.23</td>
<td>0.43</td>
<td>0.60</td>
</tr>
<tr>
<td>Change in an existing cough</td>
<td>0.47</td>
<td>0.32</td>
<td>0.15</td>
</tr>
<tr>
<td>Breathlessness</td>
<td>0.49</td>
<td>0.32</td>
<td>0.12</td>
</tr>
<tr>
<td>Discomfort in chest/shoulder/back</td>
<td>0.58</td>
<td>0.30</td>
<td>0.053</td>
</tr>
<tr>
<td>Wheezing</td>
<td>0.43</td>
<td>0.33</td>
<td>0.19</td>
</tr>
<tr>
<td>Weight</td>
<td>0.36</td>
<td>0.58</td>
<td>0.53</td>
</tr>
<tr>
<td>Recurring chest infections</td>
<td>0.03</td>
<td>0.34</td>
<td>0.93</td>
</tr>
<tr>
<td>Already seen a doctor?</td>
<td>0.59</td>
<td>0.32</td>
<td>0.06</td>
</tr>
<tr>
<td>Number of respiratory comorbidities</td>
<td>0.19</td>
<td>0.27</td>
<td>0.47</td>
</tr>
</tbody>
</table>

6.3.4 Secondary outcomes

To test for main and interaction effects of TPB-components and tailoring on the variables measured in the TPB questionnaire and the cancer risk perception measure, I conducted the Sheirer-Ray-Hare test.

As not all participants completed the secondary outcomes, I first needed to create randomly sampled equal sample sizes of questionnaire completers. As Table 6.10 shows, the maximum number of participants who completed the TPB questionnaire across all study groups was 30. I therefore randomly sampled 30 participants per study group.
Table 6.10 Number of participants who completed the TPB questionnaire per study group.

<table>
<thead>
<tr>
<th>Study group</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT</td>
<td>41</td>
</tr>
<tr>
<td>CG-TPB</td>
<td>33</td>
</tr>
<tr>
<td>CG-Tail</td>
<td>30</td>
</tr>
<tr>
<td>CG-none</td>
<td>37</td>
</tr>
</tbody>
</table>

As Table 6.11 shows, there were no significant main or interaction effects for the two experimental factors for any of the TPB-variables. This was largely confirmed in the additional four random samples (Appendix 24); across all five imputed samples I found one significant effect of tailoring for the weighted belief “Making an appointment to see a doctor about these symptoms would waste the doctor’s time”, but as this was not confirmed in the other four samples, this likely reflects random error in the data rather than a systematic effect. Similarly, I found one effect of tailoring on “If I made an appointment to see a doctor about these symptoms, the doctor would check if there is a serious cause for the symptoms”.

Table 6.11 Results of the Sheirer-Ray-Hare test for the TPB questionnaire and lung cancer risk perception scale (N=120).

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Sum of squares (SS)</th>
<th>df</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention (“I would intend to make an appointment with a doctor about these symptoms.”)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effect TPB-components</td>
<td>76.80</td>
<td>1</td>
<td>0.07</td>
<td>3.84</td>
<td>0.79</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>12.03</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>0.91</td>
</tr>
<tr>
<td>Interaction</td>
<td>403.33</td>
<td>1</td>
<td>0.39</td>
<td>3.84</td>
<td>0.53</td>
</tr>
<tr>
<td>Behavioural Attitude (“For me to make an appointment to see a doctor about these symptoms would be… [good/bad]”)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effect TPB-components</td>
<td>134.41</td>
<td>1</td>
<td>0.12</td>
<td>3.84</td>
<td>0.73</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>15.41</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>0.91</td>
</tr>
<tr>
<td>Interaction</td>
<td>145.20</td>
<td>1</td>
<td>0.13</td>
<td>3.84</td>
<td>0.72</td>
</tr>
<tr>
<td>Subjective norm (“Most people who are important to me would want me to make an appointment to see a doctor about these symptoms.”)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effect TPB-components</td>
<td>361.88</td>
<td>1</td>
<td>0.59</td>
<td>3.84</td>
<td>0.44</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>1470.01</td>
<td>1</td>
<td>2.38</td>
<td>3.84</td>
<td>0.12</td>
</tr>
<tr>
<td>Interaction</td>
<td>76.16</td>
<td>1</td>
<td>0.12</td>
<td>3.84</td>
<td>0.73</td>
</tr>
</tbody>
</table>
Composite for PBC ("For me to make an appointment to see a doctor about these symptoms would be... [difficult/easy]" and "I am confident that I could make an appointment to see a doctor about these symptoms [agree/disagree]"

<table>
<thead>
<tr>
<th>Main effect TPB-components</th>
<th>4.80</th>
<th>1</th>
<th>0.00</th>
<th>3.84</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of tailoring</td>
<td>1032.53</td>
<td>1</td>
<td>0.88</td>
<td>3.84</td>
<td>0.35</td>
</tr>
<tr>
<td>Interaction</td>
<td>252.30</td>
<td>1</td>
<td>0.21</td>
<td>3.84</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Weighted behavioural belief: "Making an appointment to see a doctor about these symptoms would waste the doctor's time."

<table>
<thead>
<tr>
<th>Main effect TPB-components</th>
<th>1274.01</th>
<th>1</th>
<th>1.08</th>
<th>3.84</th>
<th>0.30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of tailoring</td>
<td>2236.03</td>
<td>1</td>
<td>1.89</td>
<td>3.84</td>
<td>0.17</td>
</tr>
<tr>
<td>Interaction</td>
<td>118.01</td>
<td>1</td>
<td>0.10</td>
<td>3.84</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Weighted behavioural belief: "If I made an appointment to see a doctor about these symptoms, the doctor would check if there is a serious cause for the symptoms."

<table>
<thead>
<tr>
<th>Main effect TPB-components</th>
<th>790.53</th>
<th>1</th>
<th>0.67</th>
<th>3.84</th>
<th>0.41</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of tailoring</td>
<td>3488.41</td>
<td>1</td>
<td>2.97</td>
<td>3.84</td>
<td>0.09</td>
</tr>
<tr>
<td>Interaction</td>
<td>81.67</td>
<td>1</td>
<td>0.07</td>
<td>3.84</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Weighted behavioural belief: "Making an appointment to see a doctor about these symptoms, even if they are unlikely to be caused by lung cancer, can help put your mind at rest."

<table>
<thead>
<tr>
<th>Main effect TPB-components</th>
<th>6.08</th>
<th>1</th>
<th>0.01</th>
<th>3.84</th>
<th>0.94</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of tailoring</td>
<td>598.53</td>
<td>1</td>
<td>0.64</td>
<td>3.84</td>
<td>0.42</td>
</tr>
<tr>
<td>Interaction</td>
<td>3.67</td>
<td>1</td>
<td>0.00</td>
<td>3.84</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Weighted behavioural belief: "Making an appointment to see a doctor about these symptoms can help to detect lung cancer early."

<table>
<thead>
<tr>
<th>Main effect TPB-components</th>
<th>410.70</th>
<th>1</th>
<th>0.47</th>
<th>3.84</th>
<th>0.49</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of tailoring</td>
<td>452.41</td>
<td>1</td>
<td>0.52</td>
<td>3.84</td>
<td>0.47</td>
</tr>
<tr>
<td>Interaction</td>
<td>1159.41</td>
<td>1</td>
<td>1.32</td>
<td>3.84</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Weighted normative belief (family/friends): "My family and friends would want me to make an appointment to see a doctor about these symptoms"

<table>
<thead>
<tr>
<th>Main effect TPB-components</th>
<th>1373.63</th>
<th>1</th>
<th>1.16</th>
<th>3.84</th>
<th>0.28</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of tailoring</td>
<td>3520.83</td>
<td>1</td>
<td>2.98</td>
<td>3.84</td>
<td>0.08</td>
</tr>
<tr>
<td>Interaction</td>
<td>832.13</td>
<td>1</td>
<td>0.70</td>
<td>3.84</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Weighted normative belief (doctors): "My doctors would want me to make an appointment to see a doctor about these symptoms"

<table>
<thead>
<tr>
<th>Main effect TPB-components</th>
<th>869.41</th>
<th>1</th>
<th>0.75</th>
<th>3.84</th>
<th>0.39</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of tailoring</td>
<td>57.41</td>
<td>1</td>
<td>0.05</td>
<td>3.84</td>
<td>0.82</td>
</tr>
<tr>
<td>Interaction</td>
<td>120.00</td>
<td>1</td>
<td>0.10</td>
<td>3.84</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Cancer risk perception scale

<table>
<thead>
<tr>
<th>Main effect TPB-components</th>
<th>2288.13</th>
<th>1</th>
<th>1.97</th>
<th>3.84</th>
<th>0.16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of tailoring</td>
<td>7.50</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>0.94</td>
</tr>
<tr>
<td>Interaction</td>
<td>537.63</td>
<td>1</td>
<td>0.46</td>
<td>3.84</td>
<td>0.50</td>
</tr>
</tbody>
</table>
6.3.5 Affected individuals (excluding proxies)

Only 23 participants reported searching on behalf of someone else (referred to here as ‘proxies’), the remaining participants were affected individuals who reportedly experienced symptoms themselves. Proxies reported significantly higher intention than those searching for themselves (Figure 6.7), as indicated by the Mann-Whitney U-test ($U=1119.0$, $Z=-4.00$, $p<0.001$).

![Mean intention by affected individuals and proxies](image)

Figure 6.7 Mean (95% CIs) intention by affected individuals and proxies.

6.3.6 Next-steps measure

Table 6.12 shows the frequency with which the different options in the ‘Next steps’ measure were selected. There was no significant association between study group and the ‘next steps’ variable ($\chi^2(9)=11.24$, $p=0.26$).

Table 6.12 Frequency of selection of ‘Next steps’ options across the four study groups.

<table>
<thead>
<tr>
<th>Option</th>
<th>INT n=53</th>
<th>CG-TPB n=50</th>
<th>CG-Tail n=50</th>
<th>CG-none n=59</th>
</tr>
</thead>
<tbody>
<tr>
<td>Make an appointment</td>
<td>28</td>
<td>20</td>
<td>23</td>
<td>30</td>
</tr>
<tr>
<td>Register with a GP</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Change GP practice</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>None</td>
<td>25</td>
<td>28</td>
<td>26</td>
<td>29</td>
</tr>
</tbody>
</table>

6.3.7 Comparison of study groups on demographic variables

Using the Kruskal-Wallis test, I found no significant difference between study groups in age ($\chi^2(3)=1.26$, $p=0.74$), or education level ($\chi^2(3)=1.86$, $p=0.60$). Using chi-square tests, I found no
significant associations between study group allocation and gender ($\chi^2(3)=2.35$, $p=0.50$), ethnicity ($\chi^2(3)=12.67$, $p=0.39$), or smoking status [ever-smoker vs. never-smoker] ($\chi^2(3)=2.90$, $p=0.41$). Thus the four study groups did not differ significantly on any of the demographic variables I measured.

6.4 Discussion

This study aimed to test whether the presentation of information about lung cancer on the Web can be improved to enhance users’ intention to seek help. The ultimate aim was to find ways to utilise the Web to encourage people to seek help for potential lung cancer symptoms earlier, to reduce delays to diagnosis. Neither tailoring nor TPB-components nor their interaction seemed to impact on participants’ reported intention to seek help, thus not supporting Hypotheses 1 - 3. Tailoring was also not significantly associated with the proportion of participants reporting appropriate responses to the intention to seek help measure, thus not supporting Hypothesis 4.

In the following I discuss these findings and possible reasons why my hypotheses were not confirmed, and I show how I used the conclusions drawn to inform the design of a final trial.

6.4.1 TPB-components, tailoring, and intention

I found no significant main or interaction effects of the two factors, TPB-components and tailoring, on participants’ reported intention to seek help. Thus it seems neither tailoring of the lung cancer information on the website to individual user characteristics, nor the addition of components designed to target users’ beliefs about help-seeking, changed users’ reported intention to seek medical advice. It is possible that my experimental manipulation was unsuited or not strong enough to produce the hypothesised outcomes.

The TPB-components consisted of quotes by health professionals and fictional family members of patients, emphasising the benefits of early presentation of symptoms, and demonstrating that health professionals and family members endorse help-seeking. To influence control beliefs, I presented participants with a list of steps that could help to overcome barriers to making an appointment. Thus, the components involved elements of providing information about the behaviour, social encouragement, and planning. These methods are frequently used in TPB-based interventions (Steinmetz, Knappstein, Ajzen, Schmidt, & Kabst, 2016). However, a recent meta-analysis shows that TPB-based interventions tend to produce only moderate effects on
intention (Steinmetz et al., 2016), suggesting that beliefs, attitudes, perceived norms and perceptions of behaviour can be very difficult to change.

I also found no significant effects of tailoring on behavioural intention. De Nooijer et al. (2004) found effects of tailoring computer-based cancer information on intention to seek help. However, participants in De Nooijer et al.’s trial were healthy individuals responding hypothetically in relation to what they think they would do in future if they experienced a cancer symptom. In my trial, participants were advised to participate if they, or someone they knew, had any relevant symptoms. Therefore, to the best of my knowledge, participants responded in relation to actual rather than fictional symptoms. Participants’ notions in the abstract of what they might do in a fictional scenario may differ considerably from what they would do when actual symptoms arise.

A meta-analysis of Web-delivered tailored behaviour change interventions found a significant overall effect on health outcomes (Lustria et al., 2013). However, these interventions focused on habitual behaviours like nutrition/diet, alcohol consumption, smoking and physical exercise, which differ from a behaviour like help-seeking, which requires only one performance of the behaviour. Thus, it is questionable whether my intervention is comparable to previous Web-based behaviour change interventions. Furthermore, descriptions of how tailoring of information is achieved in interventions are scarce in the literature (Lustria et al., 2013), and it is thus difficult to compare my approach to information tailoring to other studies.

I did not find significant changes in the variables assessed by the TPB-questionnaire. A recent meta-analysis of 82 studies on the effects of TPB-based interventions has shown that effect sizes for changing TPB variables tend to be moderate, and effects for changes in behavioural beliefs are non-significant across studies (Steinmetz et al., 2016). This highlights the difficulty of bringing about changes in beliefs and attitudes, particularly when trying to change beliefs about the outcomes of behaviour. As the secondary outcome measures (TPB-questionnaire and risk perception item) were optional and thus not completed by all participants, these analyses may have been underpowered to detect small effects.

There are also possible methodological reasons why I may have been unable to detect effects. First, it is possible that my measurement of behavioural intention was not sufficiently sensitive. The intention measure, assessed on a 7-point scale, was heavily skewed, with 45% of
participants selecting 7 ("Yes, definitely") in response to the question “Do you intend to see a doctor about your symptoms?” Thus, it appears a ceiling effect occurred, which may have prevented us from detecting differences between the four study groups, as variation in the data beyond the top end of the scale were not reflected in the scores obtained from my measure (Po, 1998). A ceiling effect indicates that the range of the data measured by the instrument was limited, which can reduce the statistical power to detect associations with other variables (Po, 1998). It is possible that all four study groups increased participants’ intentions to seek help, or that participants had high intentions to seek help even before completing the study. I recruited mostly individuals who were in the process of researching symptoms online (using Google AdWords), and these individuals tend to be more likely than others to be worried about symptoms (White & Horvitz, 2009a, 2010b) and to seek medical help (Ybarra & Suman, 2006). This highlights that a future study should include a baseline measurement of intention prior to exposure to the information.

Another issue with the primary outcome may be its limitation as a proxy measure of actual behaviour. While intention has been shown to predict some behaviours with high accuracy (Armitage & Conner, 2001), intention is not always an accurate predictor of actual behaviour (Rhodes & Dickau, 2012; Rhodes & Plotnikoff, 2006; Webb & Sheeran, 2006). Characteristics of the behaviour under study and the context in which it is carried out play an important role; for example, intentions predict behaviour less well when there is potential for social reaction (Webb & Sheeran, 2006). For the present case – intention to seek medical help for certain symptoms – little evidence is available to estimate how well intentions may predict behaviour. Previous studies examining help-seeking for cancer symptoms tend to measure only intention, without following up on actual behaviour (de Nooijer, Lechner, & de Vries, 2003; Forster et al., 2013; Grunfeld & Kohli, 2010; Hunter et al., 2003; Morris et al., 2016; Reubsaet, van Osch, de Vries, Op de Coul, & Lechner, 2009; Trivers et al., 2011; Tustin, 2012). Osch et al. (van Osch, Lechner, Reubsaet, de Nooijer, & de Vries, 2007) found a significant association between intention to seek help for cancer symptoms in future, and past help-seeking behaviour for cancer symptoms. However, these variables were assessed cross-sectionally, raising questions about validity. Thus, I do not know whether the intention variable was a valid and accurate predictor of actual help-seeking behaviour, thereby limiting the conclusions I can draw regarding the (in-)effectiveness of the intervention.
### 6.4.2 Other factors related to intention to seek help

While intention was not significantly influenced by my experimental manipulations, it appeared to vary with age, with older participants reporting higher intentions. This is in line with previous research on help-seeking in general (Brannon, Feist, Jess, & Updegraff, 2013), and has also been documented for cancer symptoms (Whitaker, Smith, Winstanley, & Wardle, 2016) and lung cancer symptoms (Wagland et al., 2016). As the risk of lung cancer rises with age (Torre et al., 2016), higher intention to seek help among older people would seem appropriate in the context of lung cancer. However, lung cancer cases among younger people are possible, with approximately 10% of cases occurring in those aged younger than 55 (Torre et al., 2016), and efforts to raise awareness should not exclude younger populations.

Participants reporting presence of discomfort in the chest, shoulder or back reported higher intentions to seek help, though this association was not quite significant in the regression model (p=0.053). Previous research suggests that symptoms which cause higher concern and which interfere with daily life are associated with help-seeking (Whitaker et al., 2016), and pain has been identified as a facilitator to help-seeking in many studies (Safer et al., 1979; Wagland et al., 2016; Whitaker et al., 2016).

Thus, in this sample, intention to seek help was predicted by older age, and the presence of potentially alarming and interfering symptoms like discomfort/pain may also play a role.

### 6.4.3 Proxies vs. affected individuals

Those using the website on behalf of someone else were asked to report their intention to encourage the affected individual to seek help, and those who experienced symptoms themselves were asked to report intention to seek help. Proxies reported significantly higher intention than those who experienced symptoms themselves. This is in line with previous research which indicates that people are more likely to judge symptoms as serious in others than in themselves (Campbell, 1975). Moreover, a recent survey conducted by Public Health England found that adults over 50 are more likely to encourage others to seek help for respiratory symptoms than to seek help themselves (Public Health England, 2017).

This suggests a fruitful approach to encouraging help-seeking for lung cancer symptoms may be to utilise social networks around affected individuals. For example, the Be Clear on Cancer campaign encourages people to take action on symptoms they notice in others, e.g. “Has
someone you know stopped noticing their cough? A cough that has gone on for over three weeks could be a sign of lung cancer” (Department of Health, n.d.).

This finding also highlights the differences in attitudes/beliefs and behaviour between proxies and affected individuals and emphasises that these two groups should be assessed and analysed separately to allow meaningful conclusions.

6.4.4 Implications for the next trial

The discussion of my findings highlights some methodological issues which may have prevented the intervention from having stronger effects, or which may have prevented us from detecting existing effects. I therefore decided to run a further trial to address these questions. Below, I describe changes made to the trial design, how they were implemented, and the rationale for the changes.

Measuring intention pre and post treatment

I measured intention to seek help after presentation of the information, and found no significant differences between the four study groups. However, we do not know whether intention increased across all four groups – regardless of how the information was presented – or whether intentions did not change at all across the groups. This information is important to assess. Features such as information tailoring are computationally more complex than generic, static websites and also require more input from the user. If static websites can change individuals’ intentions to seek help in equal measure as more involved alternatives, this is important information for those developing health websites and public health campaigns. For the subsequent trial, I therefore introduced a baseline measure of intention to measure change in intention.

Intention measure

The intention measure showed a strong skew, indicating a ceiling effect. To mitigate this effect, I decided to present participants with more response options. Rather than measuring intention on a 7-point Likert scale, I decided to present users with a scale from 0% to 100% in steps of 10%. The psychometric literature suggests that increasing scale points up to 11 points can improve discrimination of a scale (Nunnally & Bernstein, 1994). Furthermore, a similar scale has been
used in a previous study to evaluate an intervention to increase help-seeking for lung cancer symptoms (Athey, Suckling, Tod, Walter, & Rogers, 2012).

Another issue that may have contributed to the ceiling effect could be the lack of specificity of the behaviour under study. Users were asked whether they intended to see a doctor about their symptoms. Ajzen (2006b) suggests that effects of TPB-based interventions can best be assessed when the behaviour under study is clearly defined. I therefore specified the item to “How likely do you think it is that you will see a doctor about your symptoms in the next 3 weeks?” The three week cut-off was chosen because the average waiting time for an appointment with a GP in the UK is two weeks (Pulse Today, n.d.) and it should therefore be feasible to see a doctor within this time frame. Furthermore practical considerations played a role; I followed up whether participants reported actually seeing a doctor after study participation, and a longer time frame for follow-up was not feasible given the time remaining to complete data collection.

I also added this specification to all TPB-items, e.g. “Most people who are important to me would want me to make an appointment with my doctor to have these symptoms checked” was changed to “Most people who are important to me would want me to make an appointment with my doctor within the next 3 weeks to have these symptoms checked”.

Furthermore, the discussion above indicates that there is insufficient evidence in the literature to indicate whether intention is an accurate predictor of help-seeking behaviour. Therefore, I decided to include a follow-up measure to assess whether users report actual help-seeking three weeks after completing the study.

Changes to TPB components

While the feasibility study showed some promising effects of the TPB-components, the trial showed no effect. In the feasibility study, TPB-components were presented on a separate page. During the trial, TPB-components were presented on the same page as the symptom and risk factor information. It is possible that the components were “lost” among other information.

Furthermore, my intervention presented TPB-components in a tailored format, presenting only one component to each participant depending on the item they selected in the TPB-components item. This meant that, in the INT group, users received different TPB-components and it is difficult to assess effectiveness of TPB-components as a whole. For the second trial I therefore
decided to separate tailoring and TPB-components. Rather than receiving only one TPB-component, participants in the second trial received a combination of all components, designed to target several beliefs and attitudes (the beliefs identified in Table 4.2 in Chapter 4). I decided to show this on a separate page, to ensure this information was not “lost” among symptom and risk factor information.

**Proxies vs. affected individuals**

Proxies and affected individuals differed significantly in their intentions, suggesting the intervention may have differing effects on these two groups. With a larger sample, sensitivity analyses would be feasible and useful. Due to limited resources, however, I was unable to recruit a larger sample of proxies, particularly as the feasibility study and first trial showed that only a very small proportion of users of my website are self-reported proxies. For the second trial I decided, therefore, to exclude proxies.

**Sample size calculation**

To determine the sample size for the subsequent trial, I used the pooled standard deviation calculated from Trial 1 data (SD = 2.033). I did not, however, use the means from the trial data, as I made considerable changes to the intervention and study design. Instead I assessed the sample size needed for different effect sizes. I hypothesised that the largest mean difference should be between the intervention group INT and the comparison group which involved neither tailoring nor TPB-components [CG-none], and therefore based the power calculation on this difference. Table 6.13 shows the number of participants needed for several effect sizes of varying magnitude. As suggested by Borm et al. (2007) the sample size per group was multiplied by (1−r²) to adjust for baseline measurement, where r is the expected correlation between the baseline and post-treatment measure. As I did not have an estimate for r, I used a medium correlation of 0.3 as a conservative estimate. For the subsequent trial, I aimed to recruit a minimum of 54 participants per group (N=216), in order to detect an effect size of d=0.74, which constitutes a medium effect size according to Cohen’s conventions (Cohen, 1988). A larger sample size powered to detect small effects was beyond the means available for this study.
Table 6.13 Sample sizes required to detect different effect sizes.

<table>
<thead>
<tr>
<th>Cohen’s d</th>
<th>n per group</th>
<th>Correction for baseline measurement [n^*(1-r^2) ; r=0.3]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.86</td>
<td>43</td>
<td>39</td>
</tr>
<tr>
<td>0.74</td>
<td>59</td>
<td>54</td>
</tr>
<tr>
<td>0.61</td>
<td>84</td>
<td>76</td>
</tr>
<tr>
<td>0.49</td>
<td>131</td>
<td>119</td>
</tr>
<tr>
<td>0.37</td>
<td>232</td>
<td>211</td>
</tr>
</tbody>
</table>

6.5 Conclusions

This is the first study to explore ways of presenting lung cancer information online in order to encourage earlier help-seeking. I did not find evidence to support my hypotheses that information tailoring and presentation of TPB-components, designed to target specific beliefs, would change users’ intention to seek help. However, this trial highlighted several study design issues that limit the conclusions we can draw from the findings, such as a lack of follow-up, ceiling effects on the primary outcome measure, and technical issues with the database which led to loss of data. I addressed these limitations in a further trial, which is detailed in the following chapter.
Chapter 7: Trial 2: Testing a tailored, theory-based intervention to encourage help-seeking for lung cancer symptoms: A two-way mixed factorial design

7.1 Aim

This trial (henceforth referred to as “Trial 2”) aimed to address methodological limitations identified in the trial described in Chapter 6, to better elucidate the effects of information tailoring and TPB-based components on self-reported likelihood of seeking medical help for symptoms potentially related to lung cancer, by assessing interactions with time (pre and post treatment). This chapter addresses the evaluation element of the MRC Guidance (Craig et al., 2008).

7.2 Methods

7.2.1 Intervention changes

In the original intervention described in Chapter 4, users received one tailored TPB-component to target either behavioural, normative or control beliefs, depending on how they had responded to the TPB-components item. For the present trial, I instead presented users with one generic TPB-component, which encompassed several sub-components. This generic TPB-component included quotes by health professionals and fictional family members to target behavioural and normative beliefs, as well as the list of steps to make an appointment to target control beliefs (Figure 7.1). This approach facilitated assessment of the TPB-component, as all participants received the same content rather than individually tailored components.

7.2.2 Study design

The study employed a two-way mixed factorial design, with the two binary factors ‘presence of TPB-component’ (yes/no) and ‘information tailoring’ (yes/no) (Table 7.1), and the primary outcome was assessed at two time points (pre and post treatment).

Table 7.1 Study groups for Trial 2

<table>
<thead>
<tr>
<th>TPB-component</th>
<th>Tailored information</th>
<th>Untailored information</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT</td>
<td>CG-TPB</td>
<td></td>
</tr>
<tr>
<td>CG-Tail</td>
<td>CG-none</td>
<td></td>
</tr>
</tbody>
</table>
7.2.3 Participants

*Eligibility and recruitment*

Eligibility criteria as well as recruitment strategies were the same as those used in the feasibility study (Chapter 5), except that proxies were no longer eligible for participation in trial 2.

*Sample size*

As outlined in the previous chapter, I aimed to obtain complete data for a minimum of N=216 participants. Based on Table 6.13 in Chapter 6, this should allow sufficient power to detect an effect size of d=0.74 between group INT and CG-none.

7.2.4 Outcome measures

*Primary outcome self-reported likelihood of seeing a doctor about symptoms*

Participants responded to the question “How likely do you think it is that you will see a doctor about your symptoms in the next 3 weeks?” on a scale from 0% to 100% in steps of 10% (Appendix 25), similar to the approach used in (Athey et al., 2012). This item was measured pre-treatment (before participants received information about symptoms/risk factors, TPB-components and the summary), post-treatment (after receiving information), and at follow-up if participants indicated at follow-up that their symptoms had not resolved and they had not sought help since study participation.
Figure 7.1 TPB-component used for Trial 2. Sub-components were shown to participants consecutively; participants clicked through these using the arrows and were only able to proceed with the study once all sub-components had been viewed.
Initial survey

The initial survey (measured pre-treatment) was adapted in an effort to reduce attrition. In the previous two studies, participants were required to respond to one item per symptom, to indicate whether they had experienced this symptom (yes/no). This format was relatively laborious and time-consuming to complete. For the present trial, I instead presented participants with a list of symptoms from which to select those they were experiencing. If symptoms were selected, additional questions were presented to assess duration and severity (Appendix 26). An item was also included to ask participants whether they were using the website for their own symptoms or on behalf of someone else, to identify any proxies.

‘Next steps’ measure

The ‘Next steps’ measure was modified slightly, to replace option 3 (“Change to a different GP”), which was chosen by only 2 participants in the previous trial, with the option “Find a walk-in clinic” (Appendix 27). This item was assessed post-treatment.

Optional survey

TPB questionnaire items were re-phrased to increase the specificity of the behaviour, e.g. “Most people who are important to me would want me to make an appointment with my doctor to have these symptoms checked” was changed to “Most people who are important to me would want me to make an appointment with my doctor within the next 3 weeks to have these symptoms checked” (Appendix 28). For this trial, the technical error in the database was rectified so that data were collected for the control belief and perceived power items.

The Risk Perception scale was the same as in the feasibility study (Appendix 18). The optional survey was measured post-treatment.

Follow-up survey

The follow-up survey (Appendix 29) was measured approximately 3 weeks after the first visit to the website for participants who gave an email contact and permission to follow-up. The survey asked users the following three questions, with sub-questions appearing depending on responses.

1. Whether they had contacted a doctor to make an appointment to have their symptoms checked since participating in the study
If yes, participants were asked whether they thought that the information received through the study had encouraged them to contact their doctor.

If no, proceed to question 2.

2. Whether they had discussed their symptoms with a doctor.
   - If yes, they were asked what the outcome of the consultation was (medication / referral for a chest X-ray / no treatment or referral / other).
   - If no, proceed to question 3.

3. Whether they were still experiencing the symptoms at the time of completing the follow-up survey.
   - If users reported still experiencing symptoms but not having sought help, they were asked to report their perceived likelihood of seeing a doctor about their symptoms within the next three weeks (primary outcome).
   - If no, no further questions.

**Google Analytics**

Data on website usage were collected via Google Analytics (details in Chapter 5).

7.2.5 Procedure

Ethical approval was obtained from the University Research Ethics Committee (Committee 1, Ref: 15353, approval letter in Appendix 10). The study ran between 6th April and 11th August 2017. The study procedure, shown in Figure 7.2, was largely the same as for the trial described in Chapter 6, the only difference being the introduction of a measure of self-reported likelihood of visiting a doctor to seek help at baseline and post-treatment, and the follow-up survey. If participants consenting to follow-up did not complete the follow-up survey within 10 days, one reminder email was sent. The Participant Information Sheet is shown in Appendix 30.

7.2.6 Data handling and coding

Using the same approach and criteria as in the previous trial (Table 6.2), I created a variable to denote whether responses to the post-treatment ‘self-reported likelihood of visiting a doctor’ variable were considered appropriate. Self-reported likelihood was considered high when participants scored above the mid-point of the scale (>50%) and low if they scored equal to or below the mid-point (≤ 50%).
Users enter the Homepage, receive the participant information sheet.

Informed consent

Yes

Users can leave the website if they decide not to participate.

No

Is the user ≥ 18 years and lives in the UK?

No

Users from outside the UK and <18 years are advised that they are not eligible.

Yes

Participants complete baseline questions for information tailoring (initial survey, symptoms and risk factors [Appendix 26]), and the primary outcome measure [Appendix 25].

Block randomisation in blocks of 2

INT

CG-TPB

CG-Tail

CG-none

All participants complete post-treatment battery of measures (primary outcome [Appendix 25], ‘next steps’ measure [Appendix 27], demographics [Appendix 16]).

Participant willing to complete follow-up?

Yes

Email address recorded

No

Participant willing to complete additional measures?

Yes

TPB questionnaire, risk perception scale, condition evaluation [Appendix 28, 18, 19]

No

Study end

Participants are redirected to relevant NHS information pages or the study homepage, depending on their selection in the ‘next steps’ measure.

Figure 7.2 This flowchart details the study procedure. TPB = Theory of Planned Behaviour. INT = Intervention group, CG-TPB = Comparison group 1: untailed + TPB-components, CG-Tail = Comparison group 2: tailored, no TPB-components, CG-none = Comparison group 3: untailed, no TPB-components. Changes from the previous trial are shown in green font.
7.2.7 Hypotheses

**H1:** I hypothesised a main effect of time (pre/post), with participants reporting higher likelihood of visiting a doctor after viewing the study information than before.

**H2:** I hypothesised an interaction between presence of the TPB-component and time, such that participants who received the TPB-component would show a larger increase in self-reported likelihood of visiting a doctor than those who did not.

**H3:** I hypothesised an interaction between tailoring and time, such that participants who received tailored information would show a larger increase in self-reported likelihood of visiting a doctor than those who received untailored information.

**H4:** I hypothesised an interaction between tailoring, presence of the TPB-component and time, such that participants who received a combination of the factors would show a larger increase in self-reported likelihood of visiting a doctor than those who received only one factor, or none.

**H5:** I hypothesised that the proportion of participants responding appropriately to the likelihood item post-treatment would be larger in the groups receiving tailored information than those receiving untailored information.

7.2.8 Statistical analyses

All statistical tests were carried out using a significance level of $\alpha = .05$ and using either the statistical software package IBM SPSS Statistics 22 or the software environment R (Version 3.4.1.). Due to non-normal distribution of the residuals of the primary outcome, the data violated assumptions for a parametric two-way mixed ANOVA. For non-normal data with two between-subject and one within-subject factor, Wilcox (2012) suggests running a two-way mixed ANOVA using 20% trimmed means. In this method, 20% of the largest and smallest values are removed, in order to remove the influence of data points on the tails that skew the distribution and unfairly affect the average (Wilcox, 2012). Therefore, to assess changes in self-reported likelihood of visiting a doctor from pre to post and assess main and interaction effects of the two factors (tailoring and presence of the TPB component), I conducted a robust 2-way mixed ANOVA using 20% trimmed means.
Note that theSheirer-Ray-Hare test (Holmes et al., 2016) was not used in the present study due to differential dropout from the study groups which rendered the sample sizes very unequal. As the Sheirer-Ray-Hare requires equal sample sizes, using this test would have required excluding up to 40% of participants from one group. Another option would be to compute a two-way ANOVA using M-estimators, using the change from pre to post as the outcome (Mair & Wilcox, 2016). However in this case one forces the post score to be dependent on the pre score in the linear form, whereas it might be related to a fraction of the pre score. Given these constraints, it was decided that an ANOVA using trimmed means, as suggested by Wilcox (Wilcox, 2012), is the most appropriate means of analysing the data for the present study.

To test for main and interaction effects of the two factors on secondary outcomes (TPB questionnaire and lung cancer risk perception), I conducted a two-way ANOVA using trimmed means as suggested in (Wilcox, 2012). Because the analysis of the secondary outcomes involved testing across 12 variables, a p-value adjustment using the false discovery rate (FDR) method (Benjamini & Hochberg, 1995) was used to adjust for inflation of error due to multiplicity.

For any significant effects, partial eta squared was computed as a standardised effect size measure, and assessed using the criteria 0.01, 0.06, and 0.14 for small, medium and large effects respectively (Cohen, 1988). Partial eta squared was computed using the raw (untrimmed) data, in SPSS. To further provide an estimate of the magnitude of significant effects, the unstandardised trimmed mean difference is also provided.

The Kruskal-Wallis test (Kruskal & Wallis, 1952) was used to test for differences in means across the four study groups, and Dunn’s test was used to test pairwise comparisons post hoc where significant differences across groups were found. The Mann-Whitney U-test was used to test for differences between pairs of groups. To test for associations between categorical variables (e.g. study group and smoking status), I used Pearson’s chi-squared test (Fisher, 1922; Pearson, 1992).

An exploratory sub-group analysis was conducted to test for effects of the two factors on self-reported likelihood of seeking help among those at high risk for lung cancer, i.e. those aged 50 years or above, and self-reported smokers or ex-smokers (Coureau et al., 2016).

Due to the small number of participants completing the follow-up survey, follow-up data were analysed descriptively only.
7.3 Think Aloud evaluation

Before presenting the findings of the trial, I will first present a short usability evaluation using Think Aloud methods, as considerable changes were introduced and I needed to ascertain that the altered intervention (and the study website as a whole) were easy to use and understand prior to launching the trial.

7.3.1 Participants

Five participants aged 24-56 years took part in the Think Aloud evaluation (2 female, 3 male).

7.3.2 Procedure

As in the previous Think Aloud evaluation, participants were asked to imagine they were experiencing symptoms described in a vignette. Users were instructed to click through the website and use the information to appraise the symptoms described in the vignette while verbalising their thoughts. Verbal prompts were used to elicit further information, with particular focus on the clarity, intelligibility and acceptability of the information and the website design. If users showed any signs of confusion or hesitation, they were asked to clarify why.

7.3.3 Analysis

Participants’ comments were recorded in the form of written notes during the evaluation, and subsequently grouped into broader categories.

7.3.4 Results

Participants’ comments, and the actions I undertook in consequence, are shown in Table 7.2.

Table 7.2 User feedback obtained during the Think Aloud Evaluation.

<table>
<thead>
<tr>
<th>Feedback</th>
<th>Solution/comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual design</td>
<td></td>
</tr>
<tr>
<td>The list of symptoms on the homepage looks clickable; several users tried</td>
<td>The visualisation was adapted to look more like a list with bullet points rather</td>
</tr>
<tr>
<td>to click on symptoms and appeared confused when they found they were</td>
<td>than links.</td>
</tr>
<tr>
<td>not linked.</td>
<td></td>
</tr>
<tr>
<td>Several users reported difficulty in reading some text sections with</td>
<td>The font size was increased.</td>
</tr>
<tr>
<td>smaller font.</td>
<td></td>
</tr>
<tr>
<td>Phrasing</td>
<td></td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>One user showed confusion after he had missed a question and the website would not allow him to proceed to the next page. He pointed out that the error message was not clear enough that a question had been missed.</td>
<td>Added to the error message: &quot;Looks like you missed a question. Please provide answers for all questions.&quot;</td>
</tr>
<tr>
<td>On the tailored summary page, it would be helpful to highlight that the recommendation regarding referral for a chest X-ray are based on the symptoms reported (highlight that this is tailored advice as this is not clear).</td>
<td>Added: &quot;Based on your symptoms and the 2015 NICE (National Institute for Health and Care Excellence) guidelines, the doctor should offer an urgent chest X-ray (to be performed within 2 weeks) if the symptoms persist and the cause is not known.&quot;</td>
</tr>
<tr>
<td>In the co-morbidities question (initial survey), the term “respiratory” in the answer category “other respiratory condition” was unclear to some users.</td>
<td>Added “other respiratory / lung condition”</td>
</tr>
<tr>
<td>One user pointed out that the symptom “breathlessness” would be normal under circumstances of strenuous physical exercise, and suggested it might be helpful to emphasise this.</td>
<td>Amended sentence: If you feel more out of breath than usual during activities you used to be able to do without problems (such as climbing stairs in your house/apartment, going for walks etc.), this can be a cause for concern.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Suggestions for improvement</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>One participant suggested emailing the summary page to participants</td>
<td>This change was not implemented as it was not within the scope of the factors I was testing in this study (information tailoring/theory-based content)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reading vs. skimming information</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As in the previous Think Aloud evaluation, users tended to skim over or skip risk factor information whereas symptom information was read.</td>
<td>We shortened the risk factor information to a few brief bullet points.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Positive feedback</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Several users commented positively on the generic TPB-component, saying that they liked the visualisation and the interactive nature (participants clicked through the quotes using arrows)</td>
<td>No changes required.</td>
</tr>
<tr>
<td>One user reportedly felt the website could be useful for people who struggle to communicate with GPs; this user worked in a charity organisation which was reportedly frequently accessed by people who struggled to access health services</td>
<td>No changes required.</td>
</tr>
</tbody>
</table>
Other observations

| Users reported no confusion when completing the primary outcome measure on the 11-point percentage scale | No changes required. |

7.3.5 Conclusions for the intervention/trial design

Following the Think Aloud evaluation, I implemented the changes shown in Table 7.2. Apart from minor confusion resulting from unclear phrasing or ambiguous visual design, no major issues were recorded. In terms of the new changes implemented for this particular trial, I noted that participants did not indicate any difficulty in understanding or completing the new 11-point primary outcome measure. Furthermore, participants particularly approved the new TPB-component; participants appeared to enjoy the visual design and the interaction required (participants needed to click through the quotes using arrow buttons). Therefore, after making changes to the wording to improve intelligibility and increasing font sizes, the revised website was uploaded to the server to be tested in the randomised trial. In the following, I present results of the online trial.

7.4 Results

7.4.1 Sample description

Between 6 April and 11th August 2017, 5289 users visited the website, the majority from UK-based IP addresses (5101/5289, 96.4%), according to Google Analytics data. In total, 5705 sessions took place (a session is defined as a group of user interactions, ending after 30 minutes of inactivity). Analytics data indicate that the majority of sessions began with users entering the website via paid (AdWord) searches (4301/5705) 75.4%), and the next-largest acquisition channel (752/5705, 13.2%) was via direct link-clicks (e.g. if users navigated to the website by clicking a link in an email). The remainder accessed the website from social media (300/5705, 5.3%), other websites (290/5705, 5.1%), paid advertisement on other websites via the AdWord campaign (35/5705, 0.6%), or unpaid search engine searches (27/5705, 0.5%).

The bounce rate indicates that 73.9% (approx. 3909 users) left the homepage without further interactions. In total, 1380 users were randomised (345 to INT, 344 to CG-TPB, 345 to CG-Tail, 346 to CG-none), and 270 participants completed the study (19% of those randomised,
Seventeen participants indicated they were using the website on behalf of someone else and were excluded from subsequent analyses (N=253). Sample sizes across the four groups varied considerably (Table 7.3).

**Table 7.3 Sample distribution across the four study groups.**

<table>
<thead>
<tr>
<th>Study group</th>
<th>n</th>
<th>% of total sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT (tailored, with TPB-component)</td>
<td>49</td>
<td>19.4</td>
</tr>
<tr>
<td>CG-TPB (untailored, with TPB-component)</td>
<td>58</td>
<td>22.9</td>
</tr>
<tr>
<td>CG-Tail (tailored, without TPB-component)</td>
<td>82</td>
<td>32.4</td>
</tr>
<tr>
<td>CG-none (untailored, without TPB-component)</td>
<td>64</td>
<td>25.3</td>
</tr>
</tbody>
</table>

Sample characteristics are shown in Table 7.4. The majority of participants were reportedly female (187/253, 73.9%) and of White Ethnicity (222/253, 87.7%). Approximately half had education levels below University degree level (135/253, 53.4%) and were smokers or ex-smokers (143/253, 56.5%). The most commonly reported symptom was cough (142/253, 56.1%), followed by fatigue (137/253, 54.2%), chest/shoulder/back pain (116/253, 45.8%), and breathlessness (89/253, 35.2%) (Figure 7.3), and approximately a third of all participants (81/253, 32.0%) had already presented to health services with their symptoms prior to study participation. The majority of users accessing the website via paid search engine searches had inputted keywords relating directly to lung cancer prior to clicking the study link (Table 7.5).
Figure 7.3 Symptom distribution across the sample (N=253).

Table 7.4 Self-reported sample characteristics (N=253).

<table>
<thead>
<tr>
<th>Age</th>
<th>Range: 18-86 years, M=43.1, SD=17.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>n=185 (73.1%)</td>
</tr>
<tr>
<td>Male</td>
<td>n=68 (26.9%)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>No education</td>
<td>n=6 (2.4%)</td>
</tr>
<tr>
<td>Secondary school</td>
<td>n=17 (6.7%)</td>
</tr>
<tr>
<td>Post secondary school, e.g. GCSE</td>
<td>n=38 (15.0%)</td>
</tr>
<tr>
<td>Further education, e.g. A-levels</td>
<td>n=74 (29.2%)</td>
</tr>
<tr>
<td>Undergraduate degree</td>
<td>n=61 (24.1%)</td>
</tr>
<tr>
<td>Post-graduate degree</td>
<td>n=57 (22.5%)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>n=222 (87.7%)</td>
</tr>
<tr>
<td>Black</td>
<td>n=6 (2.4%)</td>
</tr>
<tr>
<td>Asian</td>
<td>n=10 (4.0%)</td>
</tr>
<tr>
<td>Mixed</td>
<td>n=6 (2.4%)</td>
</tr>
<tr>
<td>Other</td>
<td>n=7 (2.8%)</td>
</tr>
<tr>
<td>Prefer not to say</td>
<td>n=2 (0.8%)</td>
</tr>
<tr>
<td>Smoking status</td>
<td></td>
</tr>
<tr>
<td>Never smoker</td>
<td>n=110 (43.5%)</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>n=77 (30.4%)</td>
</tr>
<tr>
<td>Current smoker</td>
<td>n=66 (26.1%)</td>
</tr>
</tbody>
</table>
Table 7.5 Key words used to access the study link via the AdWords campaign.

<table>
<thead>
<tr>
<th>Category</th>
<th>Keywords</th>
<th>Clicks (% out of 4250 clicks via Adwords campaign)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keywords relating to “lung cancer”</td>
<td>Lung cancer</td>
<td>919 (21.62%)</td>
</tr>
<tr>
<td></td>
<td>Lung cancer symptoms</td>
<td>150 (3.53%)</td>
</tr>
<tr>
<td></td>
<td>Signs of lung cancer</td>
<td>128 (3.01%)</td>
</tr>
<tr>
<td></td>
<td>Symptoms of lung cancer</td>
<td>98 (2.31%)</td>
</tr>
<tr>
<td></td>
<td>Warning signs of lung cancer</td>
<td>64 (1.51%)</td>
</tr>
<tr>
<td></td>
<td>Lung cancer signs and symptoms</td>
<td>42 (0.99%)</td>
</tr>
<tr>
<td></td>
<td>Early symptoms of lung cancer</td>
<td>25 (0.59%)</td>
</tr>
<tr>
<td></td>
<td>First signs of lung cancer</td>
<td>25 (0.59%)</td>
</tr>
<tr>
<td></td>
<td>Lung cancer warning signs</td>
<td>23 (0.54%)</td>
</tr>
<tr>
<td></td>
<td>What are the symptoms of lung cancer</td>
<td>22 (0.52%)</td>
</tr>
<tr>
<td></td>
<td>Symptoms lung cancer</td>
<td>20 (0.47%)</td>
</tr>
<tr>
<td></td>
<td>What are the signs of lung cancer</td>
<td>20 (0.47%)</td>
</tr>
<tr>
<td></td>
<td>Three symptoms of lung cancer</td>
<td>19 (0.45%)</td>
</tr>
<tr>
<td></td>
<td>Early signs of lung cancer</td>
<td>16 (0.38%)</td>
</tr>
<tr>
<td></td>
<td>Lung cancer signs</td>
<td>15 (0.35%)</td>
</tr>
<tr>
<td>Keywords relating to possible symptoms of lung cancer</td>
<td>Persistent cough</td>
<td>299 (7.04%)</td>
</tr>
<tr>
<td></td>
<td>Persistent cough lung cancer</td>
<td>29 (0.68%)</td>
</tr>
<tr>
<td></td>
<td>Chest pains</td>
<td>174 (4.09%)</td>
</tr>
<tr>
<td></td>
<td>Coughing up blood</td>
<td>160 (3.76%)</td>
</tr>
<tr>
<td></td>
<td>Pains in chest</td>
<td>23 (9.54%)</td>
</tr>
<tr>
<td></td>
<td>Pain in chest</td>
<td>17 (0.40%)</td>
</tr>
<tr>
<td></td>
<td>Hoarse voice</td>
<td>18 (0.42%)</td>
</tr>
<tr>
<td></td>
<td>Recurrent chest infections</td>
<td>17 (0.40%)</td>
</tr>
</tbody>
</table>
7.4.2 Primary outcome (self-reported likelihood of visiting a doctor)

At baseline, participants reported a mean self-reported likelihood of visiting a doctor of 43.0%, SD=38.7, and after treatment participants scored on average 50.3%, SD=38.70. Table 7.6 shows mean self-reported likelihood of visiting a doctor pre-treatment, post-treatment, and the mean difference between post and pre across the four study groups. The majority of participants indicated no change in self-reported likelihood of visiting a doctor from pre to post-treatment (165/253, 65.2%), while 78 participants (78/253, 30.8%) showed an increase and 10 (10/253, 4.0%) showed a decrease. The largest change in self-reported likelihood of visiting a doctor from pre to post was documented in INT (9.6%), the tailored group with the TPB-component, whereas the smallest change was found in CG-none (3.9%), the untailored group without TPB-components. CG-TPB and CG-Tail showed increases of 8.1% and 7.9% respectively. The Kruskal-Wallis test indicated no significant difference across the four study groups ($\chi^2 (3)=4.98, p=0.17$) (Figure 7.4).

Table 7.6 Mean self-reported likelihood of seeking help before and after treatment, and change from pre to post, across the four study groups.

<table>
<thead>
<tr>
<th>Study group</th>
<th>n</th>
<th>Pre-treatment self-reported likelihood of visiting a doctor (M, 95% CI)</th>
<th>Post-treatment self-reported likelihood of visiting a doctor (M, 95% CI)</th>
<th>Change (Post-Pre) (M, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT</td>
<td>49</td>
<td>40.20 (29.20-51.21)</td>
<td>49.80 (38.18-61.41)</td>
<td>9.59 (5.05-14.13)</td>
</tr>
<tr>
<td>CG-TPB</td>
<td>58</td>
<td>46.38 (35.37-57.39)</td>
<td>54.48 (44.03-64.93)</td>
<td>8.10 (4.12-12.09)</td>
</tr>
<tr>
<td>CG-Tail</td>
<td>82</td>
<td>43.29 (34.93-51.65)</td>
<td>51.22 (42.92-59.52)</td>
<td>7.93 (3.38-12.48)</td>
</tr>
<tr>
<td>CG-none</td>
<td>64</td>
<td>41.72 (32.35-51.09)</td>
<td>45.63 (36.13-55.12)</td>
<td>3.91 (-0.33-8.14)</td>
</tr>
</tbody>
</table>
Figure 7.4 Mean (95% CI) change in self-reported likelihood of visiting a doctor from pre to post across the four study groups.

Mean self-reported likelihood of seeking help before and after treatment across the levels of the two factors is shown in Table 7.7. Results from the two-way mixed (between-between-within) ANOVA (Table 7.8) using trimmed means showed a main effect of time, with participants reporting significantly higher likelihood of visiting a doctor to seek medical help after viewing study information as compared to before ($p<0.001$), with a large effect size of $\eta^2_\text{p} = 0.15$, thus supporting Hypothesis 1. From pre to post, self-reported likelihood of visiting a doctor increased by 11.8%. However no significant interaction between presence of the TPB-component and time (pre/post) was found ($p=0.16$), indicating that the increase in self-reported likelihood of visiting a doctor from pre to post was not affected by presence of the TPB-component and thus not supporting Hypothesis 2. Likewise, the interaction effect between tailoring and time (pre/post) was not significant ($p=0.27$), suggesting that the increase in self-reported likelihood of visiting a doctor from pre to post was not affected by tailoring, which does not support Hypothesis 3. The interaction between tailoring and the TPB-component also did not affect the difference from pre to post ($p=0.66$), thus not supporting Hypothesis 4.
Table 7.7 Mean self-reported likelihood of seeking help before and after treatment, for participants receiving tailored/untailored information, and TPB-component/no TPB-component.

<table>
<thead>
<tr>
<th></th>
<th>Pre (Mean, 95% CI)</th>
<th>Post (Mean, 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tailored</td>
<td>42.14 (35.56-48.71)</td>
<td>50.69 (44.01-57.37)</td>
</tr>
<tr>
<td>Untailored</td>
<td>43.93 (36.84-51.02)</td>
<td>49.84 (42.86-56.81)</td>
</tr>
<tr>
<td>TPB-component</td>
<td>43.55 (35.84-51.26)</td>
<td>52.34 (44.68-59.99)</td>
</tr>
<tr>
<td>No TPB-component</td>
<td>42.60 (36.44-48.77)</td>
<td>48.77 (42.58-54.96)</td>
</tr>
<tr>
<td>All</td>
<td>43.0 (38.21-47.80)</td>
<td>50.28 (45.48-55.07)</td>
</tr>
</tbody>
</table>

Table 7.8 Summary results of the two-way mixed ANOVA on 20% trimmed means.

<table>
<thead>
<tr>
<th></th>
<th>Test statistic Q</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of time (pre-post)</td>
<td>44.98</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>0.08</td>
<td>0.77</td>
</tr>
<tr>
<td>Main effect of presence of TPB component</td>
<td>0.38</td>
<td>0.54</td>
</tr>
<tr>
<td>Interaction effect tailoring * presence of TPB component</td>
<td>0.45</td>
<td>0.50</td>
</tr>
<tr>
<td>Interaction effect presence of TPB-component * time</td>
<td>1.96</td>
<td>0.16</td>
</tr>
<tr>
<td>Interaction effect tailoring * time</td>
<td>1.24</td>
<td>0.27</td>
</tr>
<tr>
<td>Interaction effect presence of TPB-components * tailoring * time</td>
<td>0.19</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Proportion of appropriate responses

Table 7.9 shows the observed and the expected count of the number of participants responding appropriately/inappropriately to the post-treatment self-reported likelihood of visiting a doctor measure, comparing tailored with untailored groups. No association between the two variables was found using Pearson’s chi-squared test ($\chi^2 (1)=1.01, p=0.32$), thus not supporting Hypothesis 5.

Table 7.9 Observed and expected count of the number of participants responding appropriately/inappropriately to post-treatment self-reported likelihood of visiting a doctor, by factor level.

<table>
<thead>
<tr>
<th></th>
<th>Responded appropriately n (expected count)</th>
<th>Responded inappropriately n (expected count)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tailored</td>
<td>63 (59.0)</td>
<td>68 (72.0)</td>
</tr>
<tr>
<td>Untailored</td>
<td>51 (55.0)</td>
<td>71 (67.0)</td>
</tr>
</tbody>
</table>
**Sensitivity analyses with high risk groups**

Because the intervention was aimed specifically at those at risk of lung cancer, I performed sensitivity analyses to explore whether those at high risk responded differently than those at lower risk. It should be noted that these analyses were exploratory and should be interpreted with caution as they required splitting the sample size and thus reducing statistical power.

**Age above 50**

For lung cancer screening trials, a cut-off of 50 years of age is often used to identify high-risk individuals (Coureau et al., 2016). I therefore analysed those aged 50 or above (n=103) separately. Table 7.10 shows the distribution of this sub-sample across the four groups.

**Table 7.10 Change (post-pre) in self-reported likelihood of visiting a doctor for those aged ≥50 years.**

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT</td>
<td>15</td>
<td>15.33 (5.55-25.12)</td>
</tr>
<tr>
<td>CG-TPB</td>
<td>18</td>
<td>5.00 (-0.19-10.19)</td>
</tr>
<tr>
<td>CG-Tail</td>
<td>42</td>
<td>5.95 (-1.40-13.30)</td>
</tr>
<tr>
<td>CG-none</td>
<td>28</td>
<td>-0.36 (-7.24-6.52)</td>
</tr>
</tbody>
</table>

The two-way mixed ANOVA with trimmed means (Table 7.11) showed a significant interaction between tailoring and time for those aged 50 or above with a small to medium effect ($p=0.013$, $\eta^2_p=0.04$). Those receiving tailored information showed a larger increase in self-reported likelihood of visiting a doctor from pre to post (13.2%) than those receiving untailored information (3.2%) (Figure 7.5).

For the interaction of the presence of the TPB-component with time there was an almost significant effect with a small effect size ($p=0.054$, $\eta^2_p=0.03$), with participants receiving the TPB-component showing a slightly larger increase in self-reported likelihood of visiting a doctor (13.8%) than those who did not receive the TPB-component (5.2%) (Figure 7.6). Thus I found some support for Hypotheses 2 and 3 among those aged ≥50 years, though in both cases effect sizes were small. It should be noted that, in both cases, confidence intervals of the means overlapped (Figure 7.5 and 7.6), indicating considerable variability in the results.
Table 7.11 Summary results of the two-way mixed ANOVA on 20% trimmed means for those aged ≥50 years.

<table>
<thead>
<tr>
<th></th>
<th>Test statistic Q</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of time (pre-post)</td>
<td>12.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>2.97</td>
<td>0.09</td>
</tr>
<tr>
<td>Main effect of presence of TPB component</td>
<td>2.21</td>
<td>0.14</td>
</tr>
<tr>
<td>Interaction effect tailoring * presence of TPB component</td>
<td>1.35</td>
<td>0.24</td>
</tr>
<tr>
<td>Interaction effect presence of TPB-component * time</td>
<td>3.73</td>
<td>0.054</td>
</tr>
<tr>
<td>Interaction effect tailoring * time</td>
<td>6.19</td>
<td>0.013</td>
</tr>
<tr>
<td>Interaction effect presence of TPB-components * tailoring * time</td>
<td>0.56</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Figure 7.5 Mean (95% CI) reported likelihood of seeking help at pre and post treatment for those who received tailored and those who received untailored information, for those aged ≥50.

Figure 7.6 Mean (95% CI) reported likelihood of seeking help at pre and post treatment for those who received the TPB-component and those who did not, for those aged ≥50.
Smokers and ex-smokers

When examining only participants who reported being smokers or ex-smokers (Table 7.12), no main or interaction effects (excepting a main effect for time) were detected with the two-way mixed ANOVA on trimmed means (Table 7.13).

Table 7.12 Change (post-pre) in self-reported likelihood of visiting a doctor, for those who were reportedly smokers or ex-smokers.

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT</td>
<td>26</td>
<td>11.53 (4.34-18.74)</td>
</tr>
<tr>
<td>CG-TPB</td>
<td>35</td>
<td>8.57 (3.16-13.98)</td>
</tr>
<tr>
<td>CG-Tail</td>
<td>43</td>
<td>10.93 (4.12-17.74)</td>
</tr>
<tr>
<td>CG-none</td>
<td>39</td>
<td>6.41 (1.67-11.14)</td>
</tr>
</tbody>
</table>

Table 7.13 Summary results of the two-way mixed ANOVA on 20% trimmed means for self-reported smokers or ex-smokers only.

<table>
<thead>
<tr>
<th>Test statistic Q</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of time (pre-post)</td>
<td>39.05725</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>0.09031634</td>
</tr>
<tr>
<td>Main effect of presence of TPB component</td>
<td>0.8183363</td>
</tr>
<tr>
<td>Interaction effect tailoring * presence of TPB component</td>
<td>0.5606061</td>
</tr>
<tr>
<td>Interaction effect presence of TPB-component * time</td>
<td>0.2845593</td>
</tr>
<tr>
<td>Interaction effect tailoring * time</td>
<td>1.733103</td>
</tr>
<tr>
<td>Interaction effect presence of TPB-components * tailoring * time</td>
<td>0.02933972</td>
</tr>
</tbody>
</table>

7.4.3 Follow-up

In all, 135 participants provided email addresses for follow-up, and 45 completed the follow-up survey (17.8% of the total sample, 45/253). Of these, 8 were in INT, 13 in CG-TPB, 13 in CG-Tail and 11 in CG-none. Twenty-five had reportedly contacted their doctor since study participation (25/45), though 5 of these indicated they had contacted a doctor for another reason other than the symptoms they had reported previously. Twenty had reportedly not contacted their doctor (20/45), and 5 had reportedly contacted their doctor, but for another reason (5/45). Twenty-two participants had reportedly discussed their symptoms with a doctor since study participation (22/45), the remainder (23/45) did not. Twenty-nine participants reported that they were still experiencing their symptoms at follow-up (29/45).
The 22 participants who had discussed their symptoms with a doctor were asked about the outcome of their consultation. Eight reported they had received a referral (8/22), 5 reported “other” (5/22), 2 reported “none” (2/22), and none reported receiving medication.

Of those who reported having contacted a doctor since study participation (n=25), 20 reported that the information received through the study had influenced their decision to contact a doctor (20/25); the remainder reported that this had not influenced their decision.

In all, eleven participants reported that they had not contacted a doctor or discussed their symptoms with a doctor since study participation, and that their symptoms were still present, and were therefore asked to report their perceived likelihood of seeking help for the symptoms within the following three weeks. Mean self-reported likelihood among these 11 participants was 21.82 (SD=25.23). There was a significant positive correlation between reported likelihood of seeking help at post-treatment, and reported help-seeking at follow-up, $r = 0.59$, $p<0.0001$.

7.4.4 Secondary outcomes

Next steps

Table 7.14 shows which actions participants indicated they planned to complete after receiving the study information. Most (175/252, 69.2%) indicated no action, but 77 participants (30.6%) indicated some form of help-seeking behaviour.

Table 7.14 Participants’ responses when asked, following treatment, what they would like to do next. Participants selected one of four options (n=252, 1 missing).

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Make an appointment</td>
<td>62</td>
<td>24.6</td>
</tr>
<tr>
<td>Register with a GP practice</td>
<td>3</td>
<td>1.2</td>
</tr>
<tr>
<td>Find a walk-in clinic</td>
<td>12</td>
<td>4.8</td>
</tr>
<tr>
<td>None</td>
<td>175</td>
<td>69.4</td>
</tr>
</tbody>
</table>

TPB questionnaire and lung cancer risk perception

As Table 7.15 shows, none of the items measured in the TPB questionnaire showed any significant differences between the four study groups. A significant p-value for the main effect of the TPB-component on perceived lung cancer risk was found, but this was rendered non-significant after FDR-adjustment.
Table 7.15 Results from the two-way ANOVA using 20% trimmed means on the secondary outcome variables.

<table>
<thead>
<tr>
<th></th>
<th>Test statistic ( Q )</th>
<th>( p )</th>
<th>( q ) (FDR-adjusted ( p )-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intention</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effect of presence of TPB-component</td>
<td>0.25</td>
<td>0.62</td>
<td>0.92</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>0.40</td>
<td>0.53</td>
<td>0.92</td>
</tr>
<tr>
<td>Interaction</td>
<td>0.10</td>
<td>0.76</td>
<td>0.92</td>
</tr>
<tr>
<td><strong>Behavioural Attitude</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effect of presence of TPB-component</td>
<td>0.02</td>
<td>0.88</td>
<td>0.92</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>0.01</td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>Interaction</td>
<td>0.07</td>
<td>0.79</td>
<td>0.92</td>
</tr>
<tr>
<td><strong>Subjective Norm</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effect of presence of TPB-component</td>
<td>1.53</td>
<td>0.22</td>
<td>0.79</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>0.86</td>
<td>0.36</td>
<td>0.92</td>
</tr>
<tr>
<td>Interaction</td>
<td>1.53</td>
<td>0.22</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>Perceived Behavioural Control</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effect of presence of TPB-component</td>
<td>0.23</td>
<td>0.64</td>
<td>0.92</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>2.69</td>
<td>0.11</td>
<td>0.50</td>
</tr>
<tr>
<td>Interaction</td>
<td>0.01</td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>Weighted behavioural belief: “Seeing a doctor about these symptoms within the next 3 weeks would waste the doctor’s time.”</td>
<td>3.23</td>
<td>0.08</td>
<td>0.50</td>
</tr>
<tr>
<td>Main effect of presence of TPB-component</td>
<td>0.04</td>
<td>0.84</td>
<td>0.92</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>2.74</td>
<td>0.10</td>
<td>0.50</td>
</tr>
<tr>
<td>Interaction</td>
<td>3.18</td>
<td>0.08</td>
<td>0.50</td>
</tr>
<tr>
<td>Weighted behavioural belief: “If I see a doctor about these symptoms within the next 3 weeks, the doctor may find a serious cause.”</td>
<td>0.04</td>
<td>0.84</td>
<td>0.92</td>
</tr>
<tr>
<td>Main effect of presence of TPB-component</td>
<td>0.4</td>
<td>0.31</td>
<td>0.86</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>1.06</td>
<td>0.31</td>
<td>0.86</td>
</tr>
<tr>
<td>Weighted behavioural belief: “Seeing a doctor about these symptoms within the next 3 weeks, even if they are unlikely to be caused by lung cancer, can help to put your mind at rest.”</td>
<td>0.65</td>
<td>0.42</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>Test statistic $Q$</td>
<td>$p$</td>
<td>$q$ (FDR-adjusted $p$-value)</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------------</td>
<td>-------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>1.71</td>
<td>0.20</td>
<td>0.79</td>
</tr>
<tr>
<td>Interaction</td>
<td>0.03</td>
<td>0.86</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Weighted behavioural belief: "Seeing a doctor about these symptoms within the next 3 weeks could help to detect lung cancer early and treat it more effectively."

<table>
<thead>
<tr>
<th></th>
<th>Test statistic $Q$</th>
<th>$p$</th>
<th>$q$ (FDR-adjusted $p$-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of presence of TPB-component</td>
<td>3.75</td>
<td>0.06</td>
<td>0.50</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>0.24</td>
<td>0.63</td>
<td>0.92</td>
</tr>
<tr>
<td>Interaction</td>
<td>1.03</td>
<td>0.31</td>
<td>0.86</td>
</tr>
</tbody>
</table>

Weighted normative beliefs (family/friends): "My family and friends would want me to see a doctor about these symptoms within the next 3 weeks."

<table>
<thead>
<tr>
<th></th>
<th>Test statistic $Q$</th>
<th>$p$</th>
<th>$q$ (FDR-adjusted $p$-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of presence of TPB-component</td>
<td>0.07</td>
<td>0.79</td>
<td>0.92</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>1.08</td>
<td>0.30</td>
<td>0.86</td>
</tr>
<tr>
<td>Interaction</td>
<td>0.09</td>
<td>0.77</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Weighted normative belief (doctors): "My doctors would want me to see a doctor about these symptoms within the next 3 weeks."

<table>
<thead>
<tr>
<th></th>
<th>Test statistic $Q$</th>
<th>$p$</th>
<th>$q$ (FDR-adjusted $p$-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of presence of TPB-component</td>
<td>0.45</td>
<td>0.50</td>
<td>0.92</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>0.62</td>
<td>0.43</td>
<td>0.92</td>
</tr>
<tr>
<td>Interaction</td>
<td>0.17</td>
<td>0.68</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Weighted control belief: "I expect that my doctor's practice will not have many appointments free in the next 3 weeks."

<table>
<thead>
<tr>
<th></th>
<th>Test statistic $Q$</th>
<th>$p$</th>
<th>$q$ (FDR-adjusted $p$-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of presence of TPB-component</td>
<td>0.33</td>
<td>0.57</td>
<td>0.92</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>2.90</td>
<td>0.09</td>
<td>0.50</td>
</tr>
<tr>
<td>Interaction</td>
<td>0.22</td>
<td>0.64</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Lung cancer risk perception

<table>
<thead>
<tr>
<th></th>
<th>Test statistic $Q$</th>
<th>$p$</th>
<th>$q$ (FDR-adjusted $p$-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effect of presence of TPB-component</td>
<td>7.17</td>
<td>0.01</td>
<td>0.36</td>
</tr>
<tr>
<td>Main effect of tailoring</td>
<td>0.45</td>
<td>0.51</td>
<td>0.92</td>
</tr>
<tr>
<td>Interaction</td>
<td>0.01</td>
<td>0.92</td>
<td>0.92</td>
</tr>
</tbody>
</table>

### 7.4.5 Group comparison on sociodemographic variables

A Kruskal-Wallis test across the four study groups indicated that the four study groups did not differ significantly in self-reported likelihood of seeking help at baseline ($\chi^2 (3)=0.45$, $p=0.93$).
Groups differed significantly in reported age ($\chi^2 (3)=9.76, p=0.02$), and Dunn’s test showed that participants in INT (M=38.63, SD=17.36) were significantly younger than participants in CG-Tail (M=46.66, SD=16.40) ($p=0.04$). The remaining pairwise comparisons were not significant ($p>0.05$). Chi-square tests did not indicate any significant differences across the four groups in terms of self-reported sex ($\chi^2 (3)=1.21, p=0.75$), smoking status ($\chi^2 (3)=5.09, p=0.53$), ethnicity [White/non-White] ($\chi^2 (3)=1.70, p=0.64$), or education level [University degree/no University degree] ($\chi^2 (3)=2.49, p=0.48$), nor in whether they had already seen a doctor about their symptoms prior to study participation ($\chi^2 (3)=6.26, p=0.10$). Using the Kruskal-Wallis test, I found no significant differences between the four groups in the number of comorbidities (the mean number of comorbidities was 0.24 for INT, 0.24 for CG-TPB, 0.32 for CG-Tail and 0.37 for CG-none), $\chi^2 (3)=1.76, p=0.62$.

7.5 Discussion

7.5.1 Main findings

This study aimed to determine whether novel forms of presenting information online (i.e. tailoring information to user characteristics and symptoms, and supplementing information with theory-based components to target specific beliefs) can increase individuals’ self-reported likelihood of visiting a doctor to seek medical help for symptoms potentially related to lung cancer. The main finding was that, across the sample of individuals aged 18-86 years, self-reported likelihood of visiting a doctor increased significantly after reading the study information (which included information about symptoms, risk factors, and when to seek medical advice), regardless of whether this information was tailored or supplemented by the TPB-component (Hypothesis 1). This suggests the current mode of presenting information about lung cancer online – generic information without components to target beliefs - may be sufficient to enhance peoples’ self-reported likelihood of seeking medical help. However, it should be noted that self-reported likelihood of seeking help remained the same in the majority of the sample (165/253). This indicates that effects detected in this study are small and limited and that further efforts are needed to improve websites about lung cancer in order to harness their potential in reducing delays to presentation.
Self-reported likelihood was significantly associated with reported help-seeking behaviour at follow-up with a large-sized correlation according to Cohen's criteria (Cohen, 1988). This provides some validation of the primary outcome measure as a proxy for help-seeking behaviour and suggests the reported findings are relevant for actual utilisation of health services.

TPB-components did not appear to affect changes in perceived likelihood of seeking help, thus not confirming Hypothesis 2. The TPB-components aimed to change beliefs about help-seeking and its outcomes, beliefs about how it is perceived by others, and control beliefs. As discussed in the previous chapter, TPB constructs can be difficult to change, and behavioural beliefs appear to be particularly resistant to change (Steinmetz, Knappstein, Ajzen, Schmidt, & Kabst, 2016). As indicated in the analysis of the secondary outcomes (i.e. the TPB questionnaire), the intervention did not appear to change any of the beliefs it aimed to target. It is difficult to ascertain whether I employed the appropriate behaviour change techniques, as previous health behaviour change interventions have mostly focused on habitual behaviours, such as exercising and diet (Steinmetz et al., 2016). Thus, further research is needed to identify strategies to target beliefs that may impede help-seeking, such as worry about wasting the doctor’s time, and fear and fatalistic beliefs about treatability.

Tailoring also appeared to have no effect on changes in participants’ self-reported likelihood of seeking help, thus not confirming Hypothesis 3. Due to resource constraints, tailoring employed in this intervention relied on relatively simple algorithms. Tailoring was based on symptoms endorsed by users and two risk factors (age and smoking), using NICE clinical guidelines (National Institute for Health and Care Excellence, 2015). It is difficult to compare my tailoring approach to that used in other studies, as articles generally provide little detail on how tailoring was achieved. Tailoring is often treated like a “unitary construct”, with interventions merely described as ‘tailored’ or ‘untailored’, but tailoring actually spans a wide range of different strategies and methods (Hawkins, Kreuter, Resnicow, Fishbein, & Dijkstra, 2008). Tailoring can involve differing degrees of customisation depending on the number, type and measurement of variables taken into account (Hawkins et al., 2008). For example, a higher degree of customisation in my intervention may yield different results, e.g. if other risk factors such as past exposure to asbestos or family history of lung cancer are taken into account. Furthermore, it is unclear whether I selected the appropriate variables to tailor messages such that attention and
perceived personal relevance are enhanced. For example, evidence suggests that people with an external locus of control respond less well to tailored health messages (Holt, Clark, Kreuter, & Scharff, 2000); thus theoretical constructs may need to be taken into account. Overall, further research is required to determine the optimum type and degree of tailoring for this context, and future studies should report their tailoring approach in detail.

When examining the sub-population of those aged 50 years and above, who are at highest risk of lung cancer (Coureau et al., 2016), a significant interaction effect of information tailoring and time (pre to post) was found, with those receiving tailored information reporting a larger increase in self-reported likelihood of seeking medical advice after receiving the information as compared to before, than those who received untailored information. Previous research has indicated that tailored information, as compared to generic information, can enhance perceived personal relevance (Kreuter et al., 2012), thereby ensuring the information is considered more carefully and has higher impact on subsequent behaviour (Petty, Barden, & Wheeler, 2009). In this study, tailoring appears to have impacted only on older users’ perceptions. Previous research has found a similar effect, with tailored websites improving attention and recall in older adults, but not in younger adults (Nguyen et al., 2017).

Many studies have indicated that older and younger Web users employ differing strategies to find and evaluate online information (e.g. Chevalier, Dommes, & Marquié, 2011; Chin & Fu, 2010; Chin et al., 2009). Specifically, research suggests that younger users are more likely to plan and regulate their online activities, whereas older users tend to rely on evaluation of search results provided by software systems like search engines (Chevalier et al., 2011). Therefore it is possible that younger users in my study tailored the information to their own needs by regulating their own activities, e.g. by skipping or skimming irrelevant information, whereas older users relied more on the tailoring function provided by the website. The effect size was small, however. The increase in self-perceived likelihood was slightly larger (13.2%) in the tailored group than in the untailored group (3.2%), though the confidence intervals overlapped, indicating uncertainty over whether the changes were meaningfully different.

Among those aged 50 years or over, I also found an almost significant effect ($p=0.054$) of TPB-components on the change in self-reported likelihood of visiting a doctor from pre to post-treatment. Participants who received the TPB-component showed a larger increase in self-
reported likelihood of seeking help from before to after viewing the study information than those who did not receive the TPB-component. The beliefs targeted by the TPB-component were derived from studies with lung cancer patients (Birt et al., 2014; Corner et al., 2005, 2006; McLachlan et al., 2015; Tod et al., 2008; Tod & Joanne, 2010) and may thus be of higher relevance for older rather than younger users. However, it should be noted that the magnitude of the difference between the two groups was again small with overlapping confidence intervals. Thus these results should be interpreted with caution and require further investigation in future research.

Hypothesis 4, which predicted an interaction between information tailoring and the TPB-component over time, was not supported by the data. The data did show a pattern in the hypothesised direction, with people who received both factors showing the largest increase in self-reported likelihood of seeking help (9.6% from pre to post), those receiving tailoring only or only the TPB-component showing slightly smaller increases (7.9% and 8.1% respectively), and participants receiving neither showing the smallest increase (3.9%). This trend was, however, not statistically significant.

The data also showed no evidence for Hypothesis 5. This hypothesis predicted that the proportion of participants responding appropriately to the likelihood item would be higher in the tailored groups than in the untailored groups, because tailored information involved a personalised recommendation of whether help should be sought. For example, in the tailored group a participant with a cough lasting more than three weeks would be advised to seek help, whereas a participant with a two-week cough would be advised to monitor their symptoms rather than make an appointment immediately. The analyses indicated no association.

However, it should be noted that the measure of ‘appropriateness’ was relatively simple and reductionist. While it was based on NICE clinical guidelines for cancer referral (National Institute for Health and Care Excellence, 2015), I relied on basic information supplied by patients and was not able to take more nuanced information into account, such as previous medical history.

To summarise, Hypothesis 1 was confirmed in this sample, and Hypothesis 4 and 5 were not confirmed. I did not find evidence to support Hypotheses 2 and 3 across the entire sample (N=253), but some evidence for Hypotheses 2 and 3 became evident when examining only older users who are at increased risk. The detected effects were small, however, and should be
assessed further prior to implementation in practice. Moreover the findings regarding older users should be interpreted with caution, as these analyses were conducted with a smaller sub-sample of n=103 and may thus be underpowered. The sub-group analyses should be regarded as exploratory and as a means to identify possible effects that may be interesting to research further in future.

7.5.2 Limitations

There are several limitations in the present study that mean that findings should be interpreted with caution.

Sample sizes across the four study groups were very uneven, with the largest number of participants in the tailored group without TPB-components (n=82) and the smallest number of participants in the tailored group with TPB-components (n=49). Differential dropout can bias study findings (Moher et al., 2010), particularly when it occurs for systematic rather than random reasons (Bell et al., 2013), e.g. when study arms differ in acceptability. As the largest number of participants remained in CG-Tail, which was the most concise condition, one might assume that the differing volume of the study information caused differential dropout. In line with this, the next largest group was in CG-Tail (n=64), which also did not include the TPB-component. Paradoxically, however, the smallest number of participants was retained in INT (n=49), although this condition, being tailored, was more concise than CG-TPB (n=58).

Unfortunately, I was unable to assess why users dropped out of the study. I attempted to implement a pop-up function to appear when users left the website to enquire after reasons, but this is not computationally possible because browsers do not allow forced redirection on exit of a website. A comparison of sociodemographic variables for those who completed the study with those who dropped out prior to completion would help to shed further light on this issue. Due to the way the website stored data, however, it was not possible to retain data for those who did not complete the study. Due to a shortage of time and resources, it was not possible to address this issue in the present study.

When comparing the four study groups in terms of sociodemographic variables, I furthermore found that participants in CG-Tail were significantly older than those in INT. This suggests that participants dropped out systematically. Younger and older adults tend to assess online health
information differently (Chevalier et al., 2011), and the findings of this study also suggest that age plays a role. Thus study findings may be biased.

Study findings are further likely to be biased by the large proportion of female participants compared to male participants. Almost three-quarters of the sample (73.1%) were women. Previous research has indicated that women are more likely than men to seek health information online (Fox & Duggan, 2013), and the difficulty of engaging men in health interventions is well documented in the literature (Jandorf et al., 2006; Malcher, 2009). It is therefore unsurprising that more female than male participants took part in this study. However, this means that results are biased towards women and may not accurately reflect men’s responses. Moreover, this indicates more needs to be done to engage men, particularly because men are at higher risk for lung cancer than women (Torre et al., 2016), and previous research has shown men’s reluctance to seek medical help (Braybrook et al., 2011; O’Mahony & Hegarty, 2009). This has been linked to masculine gender role norms (Fish, Prichard, Ettridge, Grunfeld, & Wilson, 2015). Particularly, men seem to be concerned about being perceived as a burden to others, and using under-resourced health services (O’Mahony & Hegarty, 2009). Some research findings suggest it may be possible to engage men in health interventions by using certain communication strategies such as avoiding emotive claims, or appealing to masculine belief systems (Dale, 2014). The use of such strategies, however, risks further isolating vulnerable people by perpetuating stereotypical belief systems (Dale, 2014).

Finally, it should be noted that the sample showed a higher proportion of young participants, participants with university-level education, and non-smokers than would be expected from a typical lung cancer population (Torre et al., 2016), raising concerns about generalisability to the lung cancer population. Although we incorporated strategies to engage those at high risk of lung cancer both in terms of recruitment and in terms of the website design (e.g. by targeting older people and smokers during recruitment, and ensuring the website met criteria for senior-friendliness (National Institute on Aging, n.d.)), these groups remained difficult to engage. Future endeavours to harness the Web in encouraging earlier presentation should focus on further strategies to target at-risk groups, and sufficient resources need to be allocated to these strategies.
7.6 Conclusions

This study provides evidence that it is possible to increase individuals’ self-reported likelihood of seeking help for potential lung cancer symptoms by presenting information online. Even a simple mode of presenting information in generic format (akin to lung cancer websites currently available online) can lead to significant increases in self-reported likelihood of seeing a doctor. Furthermore, this study suggests that adding a simple component to websites, which consists of images and corresponding quotes targeting beliefs about help-seeking, may, among older users, increase self-reported likelihood of seeking medical help. This component is computationally very simple, and thus could easily be incorporated by health websites. However, the detected effect size was small and further research may be required to optimise this component prior to implementation in practice. Among older users (50 years and above), it may furthermore be conducive to tailor information to specific user symptoms and risk factors. This strategy is slightly more computationally complex, however, and requires time and input from the user. Given the small effect size found in this study, the time, effort and resources required from the user and the website developer may therefore exceed the benefit gained. However, different methods of information tailoring exist (Kreuter & Skinner, 2000; Lustria et al., 2013); a different mode or degree of tailoring might produce different outcomes and thus further research should be devoted to this strategy. Overall, the effect sizes detected in this study were small and the majority of participants reported no change in self-reported likelihood of seeking help. Moreover, differential dropout from the study groups occurred, indicating that randomisation was compromised and limiting the causal conclusions we can draw from the findings. Therefore, further research is necessary before findings can be translated into practice.
Chapter 8: Overall discussion

8.1 Overview

This thesis explored the role of the Web as a health information source in the time from first symptom perception to diagnosis among people with lung cancer. I aimed to investigate whether online health information acts as a barrier or facilitator to help-seeking, and whether it can be harnessed to encourage earlier help-seeking.

Exploring the time between symptom occurrence and diagnosis is particularly important in lung cancer because previous research has revealed delays occurring during this time period which exceed recommendations (Bjerager, Palshof, Dahl, Vedsted, & Olesen, 2006; Koyi, Hillerdal, & Brandén, 2002; Olsson et al., 2009). Delays to diagnosis in cancer can lead to advanced tumour stage when treatment is commenced, rendering treatments less effective (Mackillop, 2007; Radzikowska et al., 2012). While the pathway to diagnosis (from first symptom perception to diagnosis) may be influenced by a multitude of factors relating to practitioners, patients and healthcare systems (Weller et al., 2012), in this thesis I focused on patient factors. Research has documented that symptoms are often present for several months in lung cancer patients before medical advice is sought (Corner et al., 2005, 2006), and that targeting patient factors could have considerable impacts on improving earlier diagnosis and thereby survival (Alwan et al., 2010; Hassan, 2010; C. S. Thompson & Hetzel, 1997). Several awareness-raising interventions have been launched to increase awareness about lung cancer symptoms among the public (Ironmonger et al., 2014; Power & Wardle, 2015; S. M. Smith et al., 2013), but the World Wide Web, despite being an increasingly important health information resource (Fox & Duggan, 2013), is underutilised.

Therefore, in this thesis I aimed to explore whether the Web can play a role in the time before diagnosis of lung cancer, and whether it can be used to disseminate useful information to those with relevant symptoms. At the time of conducting the research within this thesis, research on Web use for symptom appraisal was scarce and disjointed, as was highlighted in my systematic review of the literature (Mueller, Jay, Harper, Davies, et al., 2017), with very little evidence base present to guide research design and specific hypotheses. Therefore, in this work I initially adopted an exploratory approach, with the findings of each study feeding into the aims and
methods of the next study. To address the overall research aim of gaining an insight into the role of the Web in the time before a lung cancer diagnosis, I pursued the following objectives:

- To establish whether people with lung cancer access the Web to appraise their symptoms before diagnosis
- To explore whether there is potential to harness the Web as a health information source to promote earlier help-seeking
- To identify strategies for presenting online health information that can facilitate earlier help-seeking, in order to inform the development of health websites and online applications.

To meet these objectives, I initially conducted an exploratory mixed-methods study, using a brief survey among recently diagnosed lung cancer patients to identify those who had used the Web to appraise their symptoms prior to diagnosis (and those who did not), and an interview study to explore how patients, and proxies, perceived the impact of Web searches on the time before diagnosis. Drawing on insights from this initial study, in combination with findings from previous research and input from potential users, I then developed an online intervention, designed to encourage earlier help-seeking. This intervention built on psychological theory and the use of computer/Web technology to automatically present individually tailored information to users.

The intervention aimed to increase users’ intention to seek medical advice for symptoms related to lung cancer by employing two distinct techniques designed to target barriers to help-seeking which may not be addressed by lung cancer websites currently available online, or previous public health campaigns. Firstly, it tailored information to individual users to enhance perceived personal relevance, because low perceived personal relevance has limited impacts of previous lung cancer campaigns (Caswell et al., 2017). Secondly, the intervention incorporated components to address beliefs that can impede help-seeking according to previous literature.

Following development of this website, I proceeded to design a study to test its effectiveness. An initial feasibility study allowed me to test feasibility and inform sample size calculations. The two subsequent trials enabled me to test the effectiveness of the intervention in an ecologically valid setting, with the first trial highlighting questions and issues that were subsequently addressed in the second trial.
8.2 Main findings from this thesis

The first step towards achieving the aim of this work involved a systematic review of the literature on Web use for symptom appraisal (Mueller, Jay, Harper, Davies, et al., 2017). This review highlights that the Web is an important information resource in the pre-diagnosis interval, with evidence for approximately a third of the general population using the Web to appraise symptoms, and evidence indicating that online information is used to inform help-seeking decisions. It also suggests that online health information is an important resource when obtaining information from health professionals is difficult, e.g. when symptoms are embarrassing or stigmatised, when they appear trivial, or when previous visits to healthcare have been ineffective – and these scenarios have been shown to be of relevance in lung cancer (Georgios Lyratzopoulos et al., 2012; Tod et al., 2008).

Furthermore, the review indicates that there are possible issues associated with Web use for symptom appraisal that need to be considered when assessing its potential to facilitate help-seeking, such as inducing unnecessary anxiety over benign symptoms, or, conversely, complacency regarding serious symptoms. The systematic review highlights search strategies used during online symptom appraisal that helped to inform the development of the intervention, e.g. the need to elicit aspects of symptom duration and intensity from users, or the need to mitigate a ‘process of elimination’ whereby users compare their symptoms against a generic list. Finally, the literature review showed the scarcity of research on pre-diagnosis Web use among cancer populations despite its growing importance in the process of symptom appraisal and help-seeking, emphasising the important contribution of this work.

Considering the fact that Web use for health information is associated with younger age, being female, higher education level and higher income (Fox & Duggan, 2013), I expected the proportion of pre-diagnosis Web users to be low among lung cancer patients. The first study therefore set out to establish whether the Web plays a role in the time before a lung cancer diagnosis. Approximately 20% of lung cancer patients in the survey reported that either they or a proxy had used the Web to research their symptoms prior to diagnosis. This proportion is likely to increase in future, when Internet use among older adults increases, as is indicated by current trends (Zickuhr & Madden, 2012). By exploring accounts of patients and proxies during
the embedded interview study, I identified that interviewees perceived impacts of Web searches on processes in different time intervals leading up to diagnosis. Participants reportedly used online information to assess the seriousness of their symptoms and to identify possible causes (appraisal interval), to inform their decision of whether to present to health services (help-seeking interval), and to assess healthcare professionals’ advice and substantiate requests for further investigations (diagnostic interval). This helped to highlight mechanisms via which online information can facilitate earlier diagnosis, and thus informed the development of the personalised, theory-based Web intervention.

In the feasibility study, I established the practicability of recruiting participants with relevant symptoms in an online randomised controlled trial. I also obtained estimations of statistical parameters to inform sample size calculations for the subsequent trial. The feasibility study also informed changes in the study design which allowed me to obtain more meaningful insights, e.g. a factorial design allowed me to determine interaction effects. Finally, the feasibility study helped identify issues individuals may require more support with, such as communication with health professionals.

The first randomised controlled trial I conducted following feasibility assessment allowed me to test the effects of two different ways of presenting online information about lung cancer: information tailoring to individual characteristics of users; and inclusion of theory-based components, designed to target beliefs which can may as barriers to seeking help. This trial found no significant effects for either factor on reported intention to seek help. However, several methodological and technical issues arose which limited the extent to which I was able to draw meaningful conclusions from the results.

In my second trial, I addressed these limitations, and found that participants’ perceived likelihood of seeking help for their symptoms increased significantly after viewing information about symptoms and risk factors, regardless of the mode of presentation. Furthermore, when examining only those above the age of 50 years (i.e. the age category at highest risk of lung cancer (Coureau et al., 2016)), I found significant effects of information tailoring, and almost significant effects of the TPB-component, on participants’ self-reported likelihood of seeking help. The effect sizes are small, however, and differential dropout from study groups limits the conclusions that can be drawn from the findings. The majority of the sample reported no change
in perceived likelihood of seeking help, indicating that further research is necessary to explore means of harnessing the Web to encourage early presentation.

In summary, this thesis has highlighted that the Web plays a role in the pathway to diagnosis in lung cancer, and that there is potential to make use of this role in order to promote earlier diagnosis. My findings have shown that, according to perceptions of patients and proxies, online information can impact on symptom appraisal, the decision to present to health services, and on communication with doctors, including requests for referrals to specialist care. Furthermore, this thesis has provided evidence that online information - even if presented in a generic format without targeting beliefs – can significantly increase self-reported likelihood of seeking medical advice for symptoms potentially related to lung cancer. The thesis also suggests that the presentation of online information can be altered to further encourage individuals to seek help, but effect sizes were small and further research should be conducted before changes to practice are made.

Overall, this thesis has made an important contribution to research on promoting earlier presentation of lung cancer symptoms to health services, highlighting ways in which current approaches such as public health campaigns (e.g. Ironmonger et al., 2014) or primary care-based interventions (S. M. Smith et al., 2013), can be supplemented by tapping into the most widely used lay health information source, the World Wide Web.

8.3 Strengths and limitations of the studies presented within this thesis

I will now critically appraise the studies conducted within the frame of this PhD project, identifying strengths and limitations that affect the conclusions we can draw from the findings.

The systematic review (Chapter 2), which scopes out the literature around Web use for symptom appraisal, is relatively broad, including studies of any design and any symptom/disease population. This was necessary because research on this topic is relatively scarce and disjointed. However, this means that the included studies are very heterogeneous, involving quantitative and qualitative designs and various different outcome measures. A meta-analysis was not possible and thus no statistical assessment of, for example, the relationship between help-seeking and Web use, or anxiety and Web use. The broadness and diversity of included studies in the review, however, also allowed a rich, holistic understanding of the
phenomenon under study (Dixon-Woods et al., 2005). It also enabled me to scope the literature on Web use for symptom appraisal, thereby establishing that similar research had not previously been conducted in lung cancer populations and emphasising the novelty of the present work. It also enabled me to identify search strategies and information appraisal strategies commonly used during Web use for symptom appraisal, helping to inform the development of the intervention website (Chapter 4).

Having identified the lack of previous research on Web use for symptom appraisal in lung cancer, I proceeded to conduct an exploratory mixed-methods study with recently diagnosed lung cancer patients (Chapter 3). This study relied on retrospective, self-reported measures of patients, pertaining to events that took place prior to diagnosis, which can be inconsistent (Hess et al., 2012), and may include events unbeknown to the patient (e.g. Web searches conducted by proxies).

I did not statistically assess the association between the length of time intervals in the pathway to diagnosis and Web usage. Due to the small number of participants reporting Web use prior to diagnosis and the heterogeneity of the sample who reported Web use (as some were patients and some were proxies), any statistical assessment would have been underpowered in the present study. Furthermore, at the time of conducting the study, there was, to my knowledge, no fully validated measure of the intervals leading up to diagnosis available. Available measures had considerable limitations (R. Andersen et al., 2009) and were too lengthy for the purposes of the study (e.g. Neal et al., 2014). Due to the novelty of the research field, the study adopted an exploratory approach focused on qualitatively evaluating participants’ perceptions and reported experiences, rather than assessing specific constructs quantitatively. Therefore I used a brief survey to screen participants for purposive sampling for interview, rather than administering a time and resource-intensive questionnaire.

The focus on qualitative assessment means that one should be cautious when generalising from the purposive interview sample to the wider population. The aim of the interview study was not to draw generalisable conclusions about the wider population but to obtain a rich, contextual understanding (Polit & Beck, 2010) of participants’ experiences during the time leading up to diagnosis, with a focus on their perceptions of the impact of Web use for symptom appraisal. It
provided useful insights for the development of the intervention (Chapter 4) and can inform design of future research.

After obtaining an insight into lung cancer patients' (and proxies') perceptions and experiences of using online information to evaluate lung cancer symptoms, I proceeded to develop a Web-based intervention to encourage earlier help-seeking among people with relevant symptoms. A strength of the intervention was that I incorporated elements of the person-based approach (Yardley et al., 2015) by drawing on my interview findings, previous qualitative research, Patient and Public Involvement work, and Think Aloud usability studies, and by using guiding principles to ensure the intervention remained focused on its key objectives. While it was not feasible to implement the full person-based approach due to resource constraints, the authors of this approach acknowledge that this is not always practicable, and encourage researchers to implement user-centric design as far as possible (Yardley et al., 2015). A further strength of the intervention development process was the combination of the person-based approach with theory, and the use of the Behaviour Change taxonomy (Michie et al., 2008) to identify specific, observable mechanisms to bring about changes in behaviour.

Following the development of the intervention and the design of an online randomised controlled trial to test the effectiveness of the intervention, I conducted a feasibility study to test whether the proposed design was feasible in terms of recruitment and acceptable to participants, and to obtain estimates of participation levels and parameters for sample size calculations (Chapter 5). This study was underpowered to detect effects, but was not designed for this purpose. One of the main aims of this study was to identify limitations or infeasible aspects of the study design and the intervention. The feasibility study helped me to identify the shortcomings of the initial primary outcome measure ('button-click' measure), and informed the amendment of the study design into a factorial design as well as changes to the study groups to mitigate dropout.

The subsequent trial, while allowing me to assess whether information tailoring and presence of TPB-components had a significant differential effect on reported intention after presentation of the information, left several questions unanswered due to methodological and technical limitations. The trial showed that the four study groups did not differ significantly in terms of intention to seek help following exposure to the two factors, but due to a lack of baseline
measurement of intention it did not elucidate whether all groups changed intentions to the same
degree in response to the information, or not at all. Furthermore, the heterogeneity of the
sample, which included both proxies and affected individuals, limits the conclusions we can
draw from the findings. Only a small proportion of the sample were proxies (10.8%) and it was
thus not possible to analyse this group separately. Additionally, statistical power may have been
compromised due to ceiling effects in the primary outcome measure (Po, 1998).

Finally, a technical database error occurred in the feasibility study and the first trial, resulting in
loss of data for two items of the TPB questionnaire which was not noticed at the time the studies
were being conducted.

The final study addressed the majority of shortcomings identified above and thus allowed a
more detailed insight into the effects of the intervention, though some limitations to the study
should be considered.

Differential dropout occurred, indicating inclusion of participants in the study groups was not
random, which potentially biased findings. Due to technical limitations concerning how the
database stored and saved data, however, it was not possible to retain data of participants who
dropped out and thereby compare them against completers. The analyses suggest,
furthermore, that two study groups differed significantly in terms of age, and age can affect how
users evaluate online information (Chevalier et al., 2011; Chin & Fu, 2010; Chin et al., 2009).
The sample was also biased towards female participants, limiting generalisability to men, and
highlighting the need for strategies to engage men in the intervention. Future studies should
incorporate thorough analyses of the acceptability of different study groups for different sub-
groups in order to premeditate and mitigate differential dropout. This is particularly relevant for
online studies, where dropout thresholds are low (Webb et al., 2010).

It should also be noted that, while I found evidence for effectiveness of the study factors among
older participants, this sensitivity analysis involved smaller sample sizes and should therefore
be interpreted with caution. Future research should explore these potential effects further.

Furthermore, it should be noted that the questionnaire used to assess TPB variables was not a
validated measure. A standardised TPB measure across different behaviours does not exist, as
TPB questionnaires need to be adapted to the specific behaviour under study and the unique
set of beliefs that influence the behaviour (Ajzen, 2006b). The identification of relevant beliefs
was informed by previous research on barriers to help-seeking in lung cancer patients, and questionnaire structure and wording was informed by detailed guidelines provided by Ajzen (Ajzen, 2006b) as well as worked examples from the literature (Francis et al., 2004; Williams & Hamilton-West, 2013). As the questionnaire items have high face validity and as the questionnaire did not represent the primary outcome measure, I did not allocate further resources to the validation of this measure. It is possible, however, that not all relevant beliefs contributing to individuals’ decision to seek help were identified from previous literature.

Despite these limitations, this trial has several key strengths that helped elucidate effects of the two factors. Due to its mixed factorial design, the trial allowed me to compare how participants’ perceptions of their likelihood of seeking help changed from before to after viewing information, and to assess whether the degree of change differed depending on how the information was presented. Furthermore, as this study is a randomised controlled trial, which constitutes a high level of empirical evidence (Burns, Rohrich, & Chung, 2011), it is designed to allow causal conclusions rather than merely observing associations.

Another important strength of this study concerns external validity. Many studies exploring Web use for symptom appraisal and its effect on help-seeking use hypothetical symptom scenarios (Mueller, Jay, Harper, Davies, et al., 2017), instructing participants to imagine they have certain symptoms. In contrast, this study recruited participants who (reportedly) experienced actual symptoms, thus rendering the findings more applicable to the target population than findings derived from hypothetical, imagined scenarios.

Finally, the follow-up survey entails both strengths and limitations of the study. Due to a low response rate, only 17.8% of the sample was followed up. Therefore these data provide only a limited insight into participants’ help-seeking behaviour following study participation. Nevertheless, following up participants over a longer period of time allowed some validation of the primary outcome measure, by highlighting that self-reported likelihood of seeking help was significantly and positively correlated with reported actual help-seeking behaviour.

8.4 Implications

Several implications for different fields, including public health, health websites and search engine development, can be derived from the work presented within this thesis. Below, I
combine findings from the studies conducted in the frame of this thesis to portray how these could have practical impacts.

Firstly, this thesis highlights that the Web does currently play a role in some lung cancer patients’ pathway to diagnosis. At present the proportion appears to be low, but this role is likely to increase in future generations (Mueller, Jay, Harper, & Todd, 2017). If efforts to utilise this information source are only begun when use is more prevalent, we are unlikely to exploit the full potential of the Web. While possible issues associated with Web use for symptom appraisal such as elicitation of unwarranted anxiety or help-seeking remain (Mueller, Jay, Harper, Davies, et al., 2017), this thesis identifies the potential of the Web in promoting earlier diagnosis. Thus this work presents an important, early step towards recognising and incorporating the Web as a health information source that can be harnessed to complement awareness-raising efforts. This is particularly interesting because the Web represents a low-cost means of reaching large audiences, with a potential for high return on investment.

One cross-cutting theme which emerged throughout the systematic review, the interview study and the online trials was the need for further support for individuals who struggle to communicate with health professionals. The systematic review indicated that individuals can be prompted to turn to the Web when previous visits to health services have been inconclusive (Mueller, Jay, Harper, Davies, et al., 2017). In line with this, my interview study suggested that some patients and/or proxies accessed online information when they reportedly felt dissatisfied with advice received from healthcare professionals. Finally, findings from the online trials showed that approximately a third of participants had already seen a doctor about their symptoms at the time of participation, and comments suggested that some had had negative experiences. Previous awareness-raising campaigns for lung cancer have focussed on encouraging first presentation (Gordon et al., 2012; Ironmonger et al., 2014). However, interviews with lung cancer patients suggest that delays can occur when patients are left unclear on when to re-present to primary care after initial consultation (Birt et al., 2014). Overall, I therefore suggest that the public would benefit from clear information on how to proceed following first consultations, and support on how to communicate with health professionals to ensure informed decisions can be made.
Another possible implication of my findings concerns search engines and symptoms checker tools (e.g. WebMD, n.d.). The systematic review indicates that most symptom appraisal searches are symptom-based, and that search terms tend to lack specificity, e.g. in terms of symptom duration or frequency (Mueller, Jay, Harper, Davies, et al., 2017). This is corroborated by search terms reported in my survey study. Thus, strategies to ensure those searching for information on symptoms online can receive more relevant information may involve “query expansion techniques”, whereby search tools elicit further information from users in order to narrow down results. This has been shown to improve search effectiveness for diagnostic symptom searches (Palotti et al., 2015).

Findings presented in this thesis may also have implications for the design of health websites. Firstly, the survey among recently diagnosed lung cancer patients indicated that those searching for information on symptoms most frequently access general health websites with high perceived credibility, such as the NHS Choices website (NHS Choices, n.d.-i) or the WebMD website (WebMD, n.d.). This suggests that, in order to harness the Web to educate the public about lung cancer symptoms, we need to target such websites to maximise benefits. The final online randomised trial with a factorial mixed design suggests that such websites could possibly increase their impact on help-seeking behaviour for lung cancer among older users by altering how information is presented. Inclusion of simple elements, such as quotes by health professionals and fictional family members highlighting beneficial consequences and social acceptability of help-seeking, can potentially constitute useful means to encourage individuals to seek help for symptoms. The effects found in this study were small however, with large confidence intervals indicating uncertainty in the results. I would therefore suggest further research to optimise these components and increase their effectiveness prior to implementation in practice.

Tailoring information to users based on individual symptoms and risk factors may also provide a useful means of conveying information particularly to older users, who may rely more strongly on software systems to help them evaluate information than younger users (Chevalier et al., 2011). It should be noted that my findings suggest that the effect of this strategy on reported likelihood of seeking help is small, and may thus not justify the resources involved. However, different methods of information tailoring exist (Kreuter & Skinner, 2000; Lustria et al., 2013); a
different mode or degree of tailoring might produce different outcomes and thus further research should be devoted to this strategy.

Another implication for health websites may be derived from the differential attrition noted in the feasibility study and the final trial. While I was unable to assess why more users dropped out from some study groups than others, a plausible hypothesis is that users were more likely to drop out when more webpages were presented. This emphasises the importance of keeping health advice concise, and reducing the number of pages necessary to obtain relevant information.

Finally, the systematic review showed that searchers view only the top search results returned by search engines following symptom queries which appear above the “fold”, i.e. which do not require downward scrolling to become visible. To ensure this visibility, webpages about lung cancer symptoms need to ensure they are optimised for search engine algorithms, for example by ensuring they include relevant key words and titles (Google, 2010). The search terms identified in the survey among recently diagnosed lung cancer patients (Mueller, Jay, Harper, & Todd, 2017) may provide useful information for search engine optimisation.

8.5 Future research

Findings derived from this work have highlighted avenues for future research, both in terms of general Web use for symptom appraisal, and in terms of help-seeking and information-seeking behaviour among people with lung cancer (or potential lung cancer symptoms). Implications therefore concern a range of disciplines, including psychology, computer science, and public health.

Firstly, the systematic review of the literature indicates that research on Web use for symptom appraisal is relatively scarce and disjointed. Despite the broad research question and liberal eligibility criteria, only 34 studies in total were identified. It was not possible to focus the review on particular disease or symptom groups due to the wide range of populations assessed in the different studies. This suggests further research in this field, focusing on specific symptoms and conditions, is needed to enable more concerted enquiries.

The systematic review also indicates that further research is required to investigate consequences of Web use for symptom appraisal and establish directions of causality. The
review suggests links between Web use, utilisation of healthcare services, communication with health professionals, and health anxiety. However, a lack of long-term follow-up of subsequent behaviour and diagnoses and suitable comparison groups means that we cannot, given current evidence, assess whether reactions to information found online are appropriate or not. Future research should involve longitudinal assessment of Web use for symptom appraisal, following up over time whether individuals subsequently seek help or not, which diagnoses they receive, and whether help-seeking behaviour is warranted. Investigations will need to bear in mind that even in cases where a serious diagnosis is not ultimately made, initial help-seeking may nevertheless be appropriate, e.g. for a persistent cough. Research should also include appropriate comparison groups. These data can then be used to weigh the benefits of Web use for symptom appraisal (e.g., reductions in delays to diagnosis and avoidance of unnecessary health care use) against the disadvantages (e.g., unnecessary anxiety and health care use) and relate these to health care costs.

While the systematic review highlighted future research in the general field of Web use for symptom appraisal, the findings of the mixed-methods study carried implications more specifically for the pathway to diagnosis in lung cancer. The study represents an initial step towards exploring the role of the Web in the pathway to diagnosis. It was therefore designed to facilitate an open, exploratory approach. Future research should be dedicated towards more specific quantitative assessments of links between relevant variables. Using validated assessment methods, and a sufficiently powered sample, the link between Web use and the length of the pathway to diagnosis could be assessed. This could involve, for example, the detailed questionnaire developed by Neal et al. (Neal et al., 2014), though further validation is recommended, as so far only patients’ reported date of first consultation has been verified. Furthermore this should involve assessment of a range of covariates that may be related to Web use and help-seeking behaviour, such as age, gender, education and socio-economic status (Brannon et al., 2013; Fox & Duggan, 2013).

As the research presented here suggested some evidence for the effectiveness of presenting online information in different ways to encourage help-seeking, this should be explored further, in the context of lung cancer as well as other disease groups. While I found some evidence that the factors I examined changed self-reported likelihood of seeking medical help, effect sizes were small, and other measured beliefs remained unchanged. For example, the intervention did
not appear to affect users’ worry that seeking help for their symptoms may “waste the GP’s time”. This has been identified as a barrier to help-seeking in lung cancer (Tod et al., 2008; Tod & Joanne, 2010). Previous research has shown that beliefs and attitudes relevant to health behaviour are resistant to change (Steinmetz et al., 2016). Future research should therefore focus on developing strategies to target such variables, to ensure people can be encouraged to seek prompt help for lung cancer warning signs, as well as other disease groups where early diagnosis is crucial to successful treatment.

Finally, it should be emphasised that this research project focused on patient factors in the pathway to diagnosis of lung cancer. However, further research should also be dedicated to exploring factors related to health professionals and the wider healthcare system. In line with my interview findings and participants’ comments during the online study, several studies have found evidence of delays attributable to healthcare professionals and the healthcare system in the pathway to lung cancer diagnosis (Bowen & Rayner, 2002; Koyi et al., 2002; Georgios Lyratzopoulos et al., 2012). Furthermore system factors may indirectly affect patients’ delay, by influencing their beliefs about help-seeking. For example, perceived constraints of resources in the healthcare system may influence patients’ unwillingness to seek help (O’Mahony & Hegarty, 2009), and reinforce beliefs that primary care should only be accessed if a problem is extreme (Tod et al., 2008). Primary care in particular plays an increasingly important role in early diagnosis (Rubin et al., 2015). Thus, future research should seek to address issues related to the wider healthcare system and training of healthcare professionals.

8.6 Conclusion

In this thesis, I aimed to assess the role of the Web in help-seeking in lung cancer. This research topic is important because considerable delays to diagnosis in lung cancer have been documented (Corner et al., 2006), and in the UK these have been linked to low survival rates compared to the European mean (De Angelis et al., 2014; Holmberg et al., 2010). Because the largest proportion of delay has been attributed to the time between symptom occurrence and presentation to health services (Corner et al., 2005; Scott, 2014), many public health efforts have focused on raising awareness of lung cancer symptoms and encouraging the public to seek medical advice promptly for potential lung cancer symptoms, such as a persistent cough.
(e.g. Ironmonger et al., 2014). However, the World Wide Web, as the main source of lay information about health, has not been systematically assessed and utilised in this process.

Using a variety of approaches, including systematic reviewing of the literature, a survey among recently diagnosed lung cancer patients, interviews with patients and proxies, and experimental research with a randomised, factorial design, I explored the current state of the literature on Web use for symptom appraisal, whether people with lung cancer utilise the Web to appraise their condition prior to diagnosis, and finally whether changes to the way online information is presented could result in increased help-seeking for lung cancer warning signs.

Drawing on findings from these various approaches, I conclude that the Web is used by a considerable proportion of the general public to appraise symptoms and inform help-seeking decisions, and that this proportion is likely to increase in future generations. Across different disease groups, the Web is used to appraise symptoms, inform decisions of whether to seek help, and to facilitate interactions with healthcare professionals (Mueller, Jay, Harper, Davies, et al., 2017). This is mirrored in my research with recently diagnosed lung cancer patients, who described perceived impacts of Web usage on processes in the appraisal, help-seeking and diagnostic interval leading up to diagnosis. These findings demonstrate that there is potential for the Web to be utilised in the endeavour to educate the public about symptoms and to promote earlier diagnosis. Furthermore, my experimental research in ecologically valid settings – with individuals reporting actual symptoms that could be related to lung cancer – demonstrated that presentation of information about symptoms and risk factors can produce significant effects on self-reported likelihood of seeking help. Thus even simple websites (e.g. the Roy Castle Lung Cancer Foundation, n.d.) could be utilised in efforts to improve early lung cancer diagnosis and should therefore play a more prominent role in public health efforts such as the UK-wide Be Clear on Cancer campaign. Importantly, this research provides indications that alterations to health websites can change older users’ perceptions of their likelihood of seeking help, though these require further investigation. Future research needs to include strategies to ensure random allocation of participants and mitigate differential dropout to allow causal conclusions.

Overall, while I found significant effects of presenting information online, the majority of the sample reported no change in perceived likelihood of seeking medical help, indicating that further efforts are needed to fully harness the potential of the Web in promoting symptom awareness and earlier presentation to health services. The Web as a health information source
is here to stay and, if it is to be an effective tool for health care systems, websites should use evidence-based designs to help potential patients make appropriate decisions about seeking medical treatment.
References


Fitzmaurice, C., Dicker, D., Pain, A., Hamavid, H., Moradi-Lakeh, M., MacIntyre, M. F., …


249


Appendix 1: Database search strategy

(Web or internet or "search engine" or google or online or "on line") and ("help seeking" or "help-seeking" or "information seeking" or "information-seeking") and (symptom or symptoms or diagnoses or diagnosis)
### Appendix 2: Data extraction sheet

<table>
<thead>
<tr>
<th>Reference number</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Title:</td>
<td></td>
</tr>
<tr>
<td>Author:</td>
<td></td>
</tr>
<tr>
<td>Source:</td>
<td></td>
</tr>
<tr>
<td>Date:</td>
<td>Vol:</td>
</tr>
<tr>
<td>Objective:</td>
<td></td>
</tr>
<tr>
<td>Setting:</td>
<td></td>
</tr>
<tr>
<td>Study design:</td>
<td></td>
</tr>
<tr>
<td>Procedure:</td>
<td></td>
</tr>
<tr>
<td>Study population:</td>
<td></td>
</tr>
<tr>
<td>Sampling method:</td>
<td></td>
</tr>
<tr>
<td>Entry and exclusion criteria for participants:</td>
<td></td>
</tr>
<tr>
<td>Sample size:</td>
<td></td>
</tr>
<tr>
<td>Measures</td>
<td></td>
</tr>
<tr>
<td>Nature of measures:</td>
<td></td>
</tr>
<tr>
<td>Quantitative analysis</td>
<td></td>
</tr>
<tr>
<td>What statistical tests were used?:</td>
<td></td>
</tr>
<tr>
<td>Power calculation?:</td>
<td></td>
</tr>
<tr>
<td>Qualitative analysis</td>
<td></td>
</tr>
<tr>
<td>Methodology?</td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Synthesis method?</td>
<td></td>
</tr>
<tr>
<td>Relevant findings:</td>
<td></td>
</tr>
<tr>
<td>Authors’ conclusions:</td>
<td></td>
</tr>
<tr>
<td>Limitations</td>
<td></td>
</tr>
<tr>
<td>Reviewer’s conclusions</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 3: Quality appraisal of studies included in the systematic review.

Low = Poorly addressed: Reported but only vaguely; threats to the validity/reliability/objectivity of the results are probable
Medium = Addressed with minor omissions/issues: Reported but more clarification is needed (stated but details are lacking)
High = Fully addressed: Clearly reported, easy to find and understand in the text; appropriate methods to ensure high validity/reliability/objectivity

Note: For studies with mixed methods (e.g. log analyses and survey), the two methods were assessed separately to show the different methodological issues.

<table>
<thead>
<tr>
<th>Authors, date</th>
<th>Are the aims and objectives of the research clearly stated?</th>
<th>Is the research design clearly specified and appropriate for the aims and objectives of the research?</th>
<th>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</th>
<th>Do the researchers display enough data to support their interpretations and conclusions?</th>
<th>Is the method of analysis appropriate and adequately explicated?</th>
<th>Main limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attfield, Adams, &amp; Blandford, 2006</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium</td>
<td>Retrospective self-reported accounts – may be biased and/or inaccurate. No inferential/causal conclusions possible</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The aim is described succinctly in a clear sentence. The aim was to explore information seeking and use and related situational factors, and its impact on patient-provider relationships.</td>
<td>It is a qualitative design which is appropriate for the aims of exploring.</td>
<td>Study methods, measures and analysis are described, but more details on their implementation of Grounded Theory would help.</td>
<td>Identified concept are highlighted by participants’ quotes, but the authors do not state how many participants commented on each theme.</td>
<td>Grounded Theory is appropriate for the analysis of qualitative data, but for the specific questions posed here a more structured approach might have been better (e.g. Framework Analysis). Some more details on synthesis method could have been supplied.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------</td>
<td>------------------------------------------------</td>
<td>------------------------------------------------</td>
<td>------------------------------------------------</td>
<td>------------------------------------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td><strong>Briet et al., 2014</strong></td>
<td>High</td>
<td>The aim is described succinctly in a clear sentence. The aim was to explore characteristics of questions asked by users on a free health consultation website.</td>
<td>High</td>
<td>It is made clear this is a cross-sectional design, analysing the questions asked on a health website.</td>
<td>High</td>
<td>Methods are described clearly. The health website used is described, and inclusion/exclusion criteria for the questions.</td>
</tr>
<tr>
<td><strong>Cartright, White &amp; Horvitz, 2011</strong></td>
<td>High</td>
<td>The aim is described succinctly in a clear sentence. The aim was to analyse the search activity of users researching health information online, and to identify goals and patterns of search behaviour.</td>
<td>Medium</td>
<td>The study aims to examine patterns in online health searches, therefore a log-based study analysing logs from a search toolbar is appropriate. However the authors also want to explore intentions – log data without input from users cannot provide definite information on intentions.</td>
<td>Medium</td>
<td>The data used for the study is described in detail, as well as criteria for identifying different types of search sessions. But it is not clear whether/how the authors validated their algorithms.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Chin, 2009</td>
<td>High</td>
<td>High 2x2 within-subjects design was chosen to compare younger and older adults.</td>
<td>High The study procedure and measures used are described in detail. It is described how variables are operationalised and measured.</td>
<td>Medium Conclusions are based on results of ANOVA and path model, but are partly speculative, e.g. speculation that older adults use existing medical knowledge, but this was not assessed (it was only assessed whether they use an interface-driven strategy).</td>
<td>Medium ANOVA and Hierarchical multiple regression to create path model are used. ANOVA is suitable to compare groups across different conditions. The path model is suitable to create a model of search behaviour. Sample size seems small for ANOVA; power calculation is missing.</td>
<td>The sample size (n = 69) seems small for a 2x2 design. The 'older adults' group (n = 28) was considerably smaller than the younger adults group (n = 41). Quasi-experimental: We cannot draw causal conclusions about causal effects of age on search behaviour/performance. Participants are only from University community – biased sample. Experimental setting: limited external validity</td>
</tr>
<tr>
<td>Chin &amp; Fu, 2010</td>
<td>High</td>
<td>High 2x2 within-subjects design. Appropriate to compare groups and interactions between conditions.</td>
<td>High Methods are described. Study procedure and experimental conditions are described.</td>
<td>Medium Conclusions are based on statistical analyses (ANOVA) but are partly speculative, e.g. speculation that older adults use existing medical knowledge, but this was not assessed.</td>
<td>Medium ANOVA was used which is appropriate to compare groups across different conditions and analyse interactions. Sample size seems small for ANOVA; power calculation is missing.</td>
<td>The sample size (n = 46) seems small for a 2x2 design. Quasi-experimental: We cannot draw causal conclusions about causal effects of age on search behaviour/performance. Participants are only from University community – biased sample. Experimental setting: limited external validity</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Chiu, 2016</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>It was not possible to follow-up with users or assess socio-demographics; only the information provided in questions was available.</td>
</tr>
<tr>
<td></td>
<td>The aim is described succinctly in a clear sentence. The aim was to investigate questions posed to a health website.</td>
<td>Analysis of inquiries posted to a website. Suitable for the aim of describing website usage.</td>
<td>It is described how questions were selected and analysed.</td>
<td>Excerpts from users’ questions are shown to illustrate the themes.</td>
<td>Qualitative content analysis is appropriate and described in detail.</td>
<td></td>
</tr>
<tr>
<td>Cooper, Polonec, Steward, &amp; Gelb, 2013</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Retrospective and speculative self-reported accounts ('what would you do if...') – may be biased and/or inaccurate.</td>
</tr>
<tr>
<td></td>
<td>The aim is described succinctly in a clear sentence. The aim was to explore how women would evaluate symptoms associated with gynaecologic cancers.</td>
<td>Qualitative study (focus groups). Appropriate for the aim of exploring perceptions/motivations.</td>
<td>Methods and the sample are described and how focus groups were moderated and which questions were asked of participants. Qualitative analysis including development of codes is described in detail.</td>
<td>The authors illustrate how interpretations are based on data by including examples of quotes and indicating how many women endorsed which statements.</td>
<td>Qualitative analysis is appropriate for the exploratory aim and the focus group-based data. Development of codes is in line with the aims of the study. Attempts to reduce subjectivity of coding (e.g. multiple coders) were made.</td>
<td></td>
</tr>
<tr>
<td>Cumming et al., 2010</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
<td>Self-reported accounts (may be biased or inaccurate); questionnaire assesses intention which is not the same as actual behaviour.</td>
</tr>
<tr>
<td></td>
<td>The aim is described succinctly in a clear sentence. The aim was to evaluate how digital storytelling impacts on help-seeking for menopausal symptoms.</td>
<td>Cross-sectional Web-based survey. This is not appropriate to assess the impacts of the website as it did not include a control group.</td>
<td>Methods and procedure are described. More details on the survey and which questions it included would have been helpful (this can be deduced from the results section but should be explained in the methods section).</td>
<td>The authors show how interpretations relate to the data by providing excerpts of participants’ responses, and by naming percentage of participants endorsing various statements from the questionnaire.</td>
<td>Percentages of participants named. Peri- and post menopausal groups were compared using Chi square and p values are reported. Gives descriptive insight into use of the website but not its impact.</td>
<td></td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------------------------</td>
<td>----------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>De Choudhury, Ringel Morris &amp; White, 2014 LOG STUDY</td>
<td>High The aim is clearly stated in form of three clearly and succinctly formulated research questions. The aim was to research the prevalence of health activities on different platforms, i.e. social media and search engines. The authors also aimed to characterize health activities on the different platforms and describe how people evaluate information obtained from these.</td>
<td>High 15 months of log data from a major Web search engine. The log data allows the examination of search behaviour in the ‘real World’.</td>
<td>High It is described how relevant logs were identified. The study procedures and analyses of the data are described in detail.</td>
<td>High The authors base their interpretations on both participants’ survey responses and log-based data. Some issues with interpretation of log data though, e.g. all instances of health mention on Twitter were considered information ‘sharing’, though some might be ‘seeking’.</td>
<td>High Descriptive analyses, Wilcoxon test. Yes these methods are appropriate for the aim of the study, i.e. analysing the prevalence of certain search behaviours, describing the characteristics of certain search activities.</td>
<td>Twitter and logs: assumes that certain queries/tweets relate to certain topics based on certain algorithms. Some queries may be wrongly classified or the algorithms may not assess what they assume to assess.</td>
</tr>
<tr>
<td>De Choudhury, Ringel Morris &amp; White, 2014 SURVEY</td>
<td>High The aim is clearly stated in form of three clearly and succinctly formulated research questions (see above).</td>
<td>High The survey allows the researchers to validate the information obtained from logs and get deeper insights. The survey sample was selected using census representative sampling in terms of age and gender.</td>
<td>High The survey is described in detail and how it was coded and analysed.</td>
<td>Medium The authors base their interpretations on both participants’ survey responses and log-based data. Some issues with interpretation of log data though, e.g. all instances of health mention on Twitter were considered information ‘sharing’, though some might be ‘seeking’.</td>
<td>High Open coding and descriptive analyses (percentages), Wilcoxon test to test for differences between ratings of search engines and Twitter.</td>
<td>Survey: Retrospective self-reported accounts – may be biased and/or inaccurate. All survey respondents were Twitter users so perhaps biased towards younger / tech savvy users.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Fiksdal et al., 2014</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>Medium</td>
<td>Retrospective self-reported accounts – may be biased and/or inaccurate. Qualitative study, therefore limited generalisability.</td>
</tr>
<tr>
<td>Fox &amp; Duggan, 2013</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>Survey: Retrospective self-reported accounts – may be biased and/or inaccurate.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
</tr>
<tr>
<td>Hay et al., 2008</td>
<td>High The aim is described succinctly in a clearly marked sentence. The aim was to understand the extent and reasons for online research prior to first appointments for newly diagnosed rheumatology patients.</td>
<td>High Observational mixed methods study, using a survey and interviews pre and post appointments.</td>
<td>High It is described what questions were asked of patients, when, and in what setting and how patients were identified.</td>
<td>High The authors describe which tests were used to test which associations. They use descriptive statistics to describe the prevalence of online health seeking. They provide p values for associations they describe as significant. For categories from qualitative analyses, they provide quotes, and the percentage of the sample supporting these categories.</td>
<td>Medium The authors use descriptive statistics to describe the sample and prevalence of different behaviours, and correlations, chi square tests and logistic regression to test for associations. These tests are appropriate. A power analysis could have been mentioned. The qualitative analysis seems appropriate but no specific methodology was mentioned – this could have been described in more detail.</td>
<td>Survey: Retrospective self-reported accounts – may be biased and/or inaccurate.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------------------</td>
<td>------------------------------------------------------</td>
<td>------------------------------------------------------</td>
<td>------------------------------------------------------</td>
<td>------------------------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Keselman, Browne &amp; Kaufman, 2008</td>
<td>High</td>
<td>The aim is described in two clear succinct sentences. The aim was to explore users' information seeking difficulties by conceptualizing information seeking as a form of hypothesis testing, and to examine the role of users competencies in online information seeking.</td>
<td>High Qualitative interview and Think Aloud study. Appropriate for the aim of exploring phenomenon and for assessing Web use behaviour.</td>
<td>Medium Methods are described and it is clear how participants were interviewed, which tasks they completed and how this was assessed (Think Aloud). Participants’ ages are not mentioned!</td>
<td>High The authors cluster participants into 3 clusters based on their search behaviour and then describe in detail the behaviour and experiences of participants in each cluster; they provide examples and quotes for each cluster and the number of participants in each cluster.</td>
<td>Medium The authors use semantic analysis and thematic analysis which is suitable for qualitative data. The process of qualitative analysis (e.g. how were codes developed?) could have been described in more detail.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Lauckner &amp; Hsieh, 2013</td>
<td>High The aim is described clearly and succinctly in form of 4 hypotheses and 2 research questions. Does the position and frequency of serious conditions in search results affect perceived severity and susceptibility, and is this related to negative emotional outcomes? Do health literacy and experience with online health seeking moderate these relationships?</td>
<td>High 2x2 within-subjects design. Yes this is appropriate to test the research hypotheses, to test for interaction and moderation effects. The experimental design means it is appropriate to test for causal links which is part of the aims.</td>
<td>High The experimental conditions are described in detail as well as the measures used to operationalise the variables of interest.</td>
<td>High For each hypothesis and research question the authors describe the strength of associations (correlations), and significance values for main effects and interactions.</td>
<td>Medium The authors use ANCOVA which is suitable to test for main effects and interactions, and path analysis to explore the causal chain of effects. A power calculation should have been mentioned.</td>
<td>Sample consists only of University undergraduates – biased. Experimental setting: limited external validity.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Luger, 2014</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>All participants had university education, 52% postgrads – biased sample. Experimental setting: limited external validity</td>
</tr>
<tr>
<td></td>
<td>The aim is described in a clear and succinct sentence. The aim was to explore older adults’ online health seeking to determine the cognitive and diagnostic processes involved.</td>
<td>2x2 within-subjects, mixed methods design (embedded qualitative Think Aloud element). It is not clarified why they chose to vary the two factors (search tool and symptom vignette) and this is not incorporated into the analysis.</td>
<td>The conditions are described in detail, as well as recruitment strategies, eligibility criteria, and measures used to assess outcome measures.</td>
<td>The authors use qualitative analysis of Think Aloud to demonstrate participants’ search behaviours. The authors report the percentage of participants engaging in different behaviours (p values and correlations are described for quantitative analysis but these results are not relevant for the review and are thus not reviewed here).</td>
<td>The Think Aloud method is appropriate to explore participants’ search behaviour. The authors could have described their qualitative methodology and synthesis method in more detail.</td>
<td></td>
</tr>
<tr>
<td>Medlock et al., 2015</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>High</td>
<td>High</td>
<td>Survey: Retrospective self-reported accounts – may be biased and/or inaccurate. Convenience sampling via a Christian senior’s organisation.</td>
</tr>
<tr>
<td></td>
<td>The aim is described in a clear and succinct sentence. The aim was info sources seniors who use the Internet use for health information.</td>
<td>Online survey with convenience sampling via a Christian seniors organisation. The survey method is appropriate for the aims, but the sampling method may have resulted in bias.</td>
<td>Methods are described, however not eligibility criteria for participants are mentioned; perhaps an age cut-off would have been useful to focus on seniors?</td>
<td>The authors describe the sample based on responses to the questionnaire.</td>
<td>Linear regression was used to assess associations between variables. False discovery rate used to adjust p-values for multiple hypothesis testing (good methods – this is often not taken into account).</td>
<td></td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Morgan et al., 2014</td>
<td>High The aim is described in a clear and succinct sentence: The aim was to describe what information people seek from a US website about genetic and rare diseases, and why, and compare this across two years.</td>
<td>High Longitudinal, observational analysis of inquiries posted to a website. Suitable for the aim of describing website usage.</td>
<td>High The website from which the inquiries were obtained is described in detail, as well as criteria used to choose which inquiries to include, the sample of inquiries (number etc.) and which data was available for users posting the inquiries.</td>
<td>High The authors use descriptive statistics to describe the users, and name the percentages of inquiries relating to specific topics/questions. Inferential statistics are used to compare differences between inquiries of the two years (2006 and 2011)</td>
<td>High The authors use content analysis to analyse the questions, which is appropriate for their aim of systematically describing the written content on the website.</td>
<td>Sample in 2006 (n = 68) is considerably smaller than in 2011 (n = 210). Specific to genetic diseases.</td>
</tr>
<tr>
<td>Norr, Capron &amp; Schmidt, 2014</td>
<td>High The aim was to investigate whether viewing medical websites may adversely affect anxiety sensitivity (AS), and whether this is moderated by intolerance of uncertainty.</td>
<td>Medium Experimental design; participants were randomised to view either websites about symptoms or general health websites. Appropriate to determine causal effects. No power calculation mentioned. A convenience sample of undergraduate students is used so potential bias (younger, female, educated)</td>
<td>High Participants, measures, procedures and experimental manipulations are described.</td>
<td>High The authors clearly base their conclusions on the findings of the study, and are justified in drawing causal conclusions due to the experimental design.</td>
<td>High The authors use chi square and t tests to test whether participants in the experimental and control conditions differed. Hierarchical regression to assess whether the experimental manipulation predicts AS.</td>
<td>Experimental setting: limited external validity.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>North et al., 2011</td>
<td>High The aim is described in a succinct clear sentence: The aim was to establish what symptoms internet users tend to look up online, and whether this differs from telephone triage use.</td>
<td>High Cross-sectional analysis of clicks on a health website and calls to a telephone triage system. Suitable for comparison.</td>
<td>High It is described which data were included in analysis (clicks to specific symptoms on the Mayo symptom checker website, and calls to Ask Mayo Clinic) and how data were categorized for analysis.</td>
<td>High The authors report correlations between calls and click relating to specific symptoms as well as Odds Ratios, which illustrates how use of a website differs from use of a telephone triage system. The authors report the strength of associations (correlations) and significance values to support their interpretations.</td>
<td>High Odds Ratio are suitable to determine which symptoms are researched more on the Internet vs. telephone, and correlations are suitable to show whether there are relationships between use of internet and telephone for specific symptoms.</td>
<td>Conclusions about users' motivations were based on analysis of clicks rather than asking users directly. This may not reflect true motivation.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------------------</td>
<td>----------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Perez et al., 2015</td>
<td>High</td>
<td>Described clearly. The aim was to describe Internet search processes, and identify demographic and personal characteristics associated with use of System 1 and System 2 processing (system 2: including a hypothesis testing and evidence gathering, system 1: all others)</td>
<td>Medium</td>
<td>High</td>
<td>It is described how participants were recruited, eligibility criteria, study procedure, how data were collected and coded. Problematic: “Unless the participant inquired, participants were not informed that the symptoms were suggestive of influenza or meningitis.” Does this mean participants were told if they asked during the experiment?? That would reduce the validity of findings.</td>
<td>Medium</td>
</tr>
<tr>
<td>Powell et al., 2011</td>
<td>High</td>
<td>The aims are described in a clear, succinct sentence. The aims was to identify the characteristics and motivations of online health information seekers accessing the NHS Direct website.</td>
<td>High</td>
<td>Low</td>
<td>The interviews are described, though more detail could have been provided on the analysis. The survey measure is only described very vaguely, no details provided.</td>
<td>High</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Powley et al., 2016</td>
<td>High&lt;br&gt;The aims are described in a clear, succinct sentence. The aims was to assess whether patients with inflammatory arthritis use the Internet to appraise their symptoms, and to assess outcomes of symptom checker tools when used by patients.</td>
<td>High&lt;br&gt;A survey and observational study (observation of patients using symptom checker tools) is suitable to evaluate Internet use and outcomes of symptom checker tools.</td>
<td>Medium&lt;br&gt;Mostly clear, but it is not stated whether patients completed the study before or after obtaining their diagnosis. This might affect how they complete the symptom checker tool. It is not clearly stated how patients’ searches and outcomes were recorded (video? audio? screenshots?)</td>
<td>High&lt;br&gt;The authors show the number and proportion of patients reporting Internet use, which information patients inputted into the symptom checker tools, and which advice and diagnoses are suggested by the tools.</td>
<td>Low&lt;br&gt;The authors mostly just report counts and proportions. But the authors use some inferential statistics though their sample size is very small (N = 34) and likely underpowered; no power analysis is mentioned.</td>
<td>Small sample size; likely underpowered for inferential statistics. Retrospective self-reported accounts – may be biased and/or inaccurate. For the observational part (observing patients completing a symptom checker tool), validity is questionable – Would users in real settings rely on just symptom checkers?</td>
</tr>
<tr>
<td>Rice, 2006</td>
<td>High&lt;br&gt;The aims are described in a clear, succinct sentence. The aim was to understand what influences online health seeking, what the reported benefits of online health seeking are, and to identify similarities among online activities.</td>
<td>Medium&lt;br&gt;Cross-sectional survey design; analysis of existing dataset from Pew Internet and American Life project. This is suitable for description but not to identify influencing factors.</td>
<td>Medium&lt;br&gt;The datasets are described as well as which variables were included. Methods and results appear to be mixed together so it is difficult to ascertain analysis methods used.</td>
<td>High&lt;br&gt;The authors draw conclusions about which variables are associated with online health seeking and provide significance values as well as R square and standardised beta coefficients for their models.</td>
<td>High&lt;br&gt;The authors use cross-tabulations, logistic regressions, and multidimensional scaling. These methods are appropriate to establish predictors of online health seeking and to identify similarities between Internet activities.</td>
<td>Limited relevance as they do not distinguish between those who search pre and post diagnosis. Retrospective self-reported accounts – may be biased and/or inaccurate. Only asked about using the Internet to self-diagnose without seeing a doctor, which may exclude those who researched online and then saw a doctor.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Simons et al., 2017</td>
<td>High Yes, the aim was to explore peoples’ response to symptoms of rheumatoid arthritis, compared to other illnesses which are better publicised</td>
<td>Medium Participants describes hypothetical responses to symptom vignettes; this provides limited insight into actual responses to symptoms</td>
<td>High The vignettes and interview procedure are described as well as the thematic analysis process</td>
<td>High All themes are illustrated using interview quotes</td>
<td>High Thematic analysis is appropriate for qualitative exploration with specific research questions</td>
<td>Experimental setting: limited external validity.</td>
</tr>
<tr>
<td>Teriaky, Tangri &amp; Chande, 2015</td>
<td>High The purpose of this study is to understand how outpatients awaiting initial gastroenterology consultation seek medical information on the Internet.</td>
<td>Medium Cross-sectional survey with patients awaiting consultation at a gastroenterology clinic. Appropriate to describe how individuals used the internet. Sample size seems small (N = 87) and no power calculation reported.</td>
<td>High Design, procedure, material and methods are described in sufficient detail.</td>
<td>High Conclusions are based on the associations reported in the results.</td>
<td>High The authors use descriptive statistics, and chi square / t tests to test differences between those who used the Internet and those who did not.</td>
<td>Retrospective self-reported accounts – may be biased and/or inaccurate.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Thomson, Siminoff &amp; Longo, 2012</td>
<td>High The aims is described in a clear succinct sentence: The aim was to explore characteristics of colorectal cancer patients who used the Web to appraise symptoms prior to diagnosis.</td>
<td>High Cross sectional survey design. This design is appropriate for the aim of establishing characteristics of those using the Web for symptom appraisal and describing their experiences.</td>
<td>High It is described how patients were identified, eligibility criteria, and which questions were asked during interviews and how these were conducted.</td>
<td>Medium The authors provide correlations and significance values for the factors they report as associated with Web use. However they conclude WUPD did not ‘influence’ patient delay. This study design does not allow causality inferences.</td>
<td>High It is described how interview data was coded to allow quantitative analysis. The authors use t-tests, chi-square tests, ANOVA, logistic &amp; linear regression, and correlation analysis. These methods are appropriate to determine the characteristics of those who use the Web. The sample size seems sufficiently large.</td>
<td>Retrospective self-reported accounts – may be biased and/or inaccurate.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>White &amp; Horvitz, 2009</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>Low</td>
<td>High</td>
<td>Log based study – assumes that certain queries relate to certain topics based on certain algorithms. Some queries may therefore be wrongly classified or the algorithms may not assess what they assume to assess.</td>
</tr>
<tr>
<td>LOG STUDY</td>
<td>The aims are described in form of 2 clear research objectives: 1) to describe escalations that occur when users search for common symptoms and this escalates to queries about serious conditions, and 2) how this persists over several sessions.</td>
<td>Log-based longitudinal study (using logs from windows live toolbar). Log data can provide information about behaviours of ‘real’ users (but limited as it is based on assumptions).</td>
<td>It is described in detail how log data was obtained, how it was analysed to identify users who were searching for symptoms, and how their subsequent behaviour (escalation). But some numbers are only vaguely reported, e.g. the number of users from whom log data was collected (‘hundreds of thousands’).</td>
<td>The authors describe how they identify escalations and interruptions and how this can interrupt subsequent searches. But the authors tend to overstate, concluding causal relationships from cross-sectional observational data and equating ‘escalations’ with increased anxiety. The validity of interruptions was tested by visual inspection by one of the authors – but even visual inspection cannot verify searchers’ true intentions. The validity of other assumptions was not tested.</td>
<td>The authors provide descriptive statistics to describe user behaviour (percentages of logs). Tukey’s post-hoc tests, independent measures t test, ANOVA, multiple regression analysis. These tests are appropriate to describe whether searches containing escalations differ from those without escalations.</td>
<td></td>
</tr>
</tbody>
</table>

White & Horvitz, 2009

LOG STUDY

The aims are described in form of 2 clear research objectives: 1) to describe escalations that occur when users search for common symptoms and this escalates to queries about serious conditions, and 2) how this persists over several sessions.

Log-based longitudinal study (using logs from windows live toolbar). Log data can provide information about behaviours of ‘real’ users (but limited as it is based on assumptions).

It is described in detail how log data was obtained, how it was analysed to identify users who were searching for symptoms, and how their subsequent behaviour (escalation). But some numbers are only vaguely reported, e.g. the number of users from whom log data was collected (‘hundreds of thousands’).

The authors describe how they identify escalations and interruptions and how this can interrupt subsequent searches. But the authors tend to overstate, concluding causal relationships from cross-sectional observational data and equating ‘escalations’ with increased anxiety. The validity of interruptions was tested by visual inspection by one of the authors – but even visual inspection cannot verify searchers’ true intentions. The validity of other assumptions was not tested.

The authors provide descriptive statistics to describe user behaviour (percentages of logs). Tukey’s post-hoc tests, independent measures t test, ANOVA, multiple regression analysis. These tests are appropriate to describe whether searches containing escalations differ from those without escalations.

Log based study – assumes that certain queries relate to certain topics based on certain algorithms. Some queries may therefore be wrongly classified or the algorithms may not assess what they assume to assess.
<table>
<thead>
<tr>
<th>Authors, date</th>
<th>Are the aims and objectives of the research clearly stated?</th>
<th>Is the research design clearly specified and appropriate for the aims and objectives of the research?</th>
<th>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</th>
<th>Do the researchers display enough data to support their interpretations and conclusions?</th>
<th>Is the method of analysis appropriate and adequately explicated?</th>
<th>Main limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>White &amp; Horvitz, 2009 a</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>The survey sample is heavily biased towards young, educated male respondents. Retrospective self-reported accounts – may be biased and/or inaccurate.</td>
</tr>
<tr>
<td>SURVEY</td>
<td>The aims are described in form of 2 clear research objectives: 1) to describe escalations that occur when users search for common symptoms and this escalates to queries about serious conditions, and 2) how this persists over several sessions.</td>
<td>Cross-sectional survey. Survey is used to supplement and validate log data. This is appropriate but the log sample and survey sample are likely significantly different, as the survey sample is heavily biased (Microsoft employees).</td>
<td>The survey is described, i.e. which questions are asked and how participants were recruited.</td>
<td>The authors describe the survey results. But the authors tend to overstate, concluding causal relationships from cross-sectional observational data. Furthermore they identified hypochondria through two simple self-reported items; a validated scale would have been more appropriate.</td>
<td>The authors provide descriptive statistics for the survey data.</td>
<td></td>
</tr>
<tr>
<td>White &amp; Horvitz, 2009 b</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>Retrospective self-reported accounts – may be biased and/or inaccurate. The survey sample is heavily biased towards young, educated male respondents.</td>
</tr>
<tr>
<td></td>
<td>The aims are described in a clear and succinct sentence: The aim was to explore how lay individuals use the Web to find explanations for symptoms, what activities they pursue and what their experiences are.</td>
<td>Cross-sectional survey study. This is appropriate for exploring and describing behaviour and experiences of individuals. But the sample is heavily biased (Microsoft employees).</td>
<td>The survey described in detail, i.e. which questions were asked, how it was administered, how participants were identified and recruited.</td>
<td>The authors provide percentages when describing responses, and significance values when describing differences between respondent groups. Survey results are partly overstated, i.e. causal conclusions based on observational, cross-sectional data (the Web influences anxiety levels)</td>
<td>The authors use chi square test and independent t tests to test for significance of observed differences; appropriate test to determine differences between Web users, e.g. between those who report themselves as hypochondriacs compared to the rest.</td>
<td></td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>White &amp; Horvitz, 2010 a</td>
<td>High The aim is described in a clear and succinct sentence: The aim was to predict escalations in searches based on characteristics of websites visited.</td>
<td>High Cross sectional log study using logs from windows live toolbar. This is an appropriate design for analysing and describing escalations based on Web site characteristics.</td>
<td>Medium The authors describe in detail how they identified and defined escalations, and which traits of websites they examined. But the exact number of logs mined for the study is not given.</td>
<td>High The authors demonstrate the predictive value of their model and which website traits are significantly related to escalations.</td>
<td>Medium The authors use logistic regression to establish a model predicting escalations based on website traits, however beta coefficients and R square values are not reported.</td>
<td>Log based study – assumes that certain queries relate to certain topics based on certain algorithms. Some queries may therefore be wrongly classified or the algorithms may not assess what they assume to assess.</td>
</tr>
<tr>
<td>White &amp; Horvitz, 2010 b</td>
<td>High The aim is described in a clear succinct sentence: The aim was to establish predictors of when searches for common symptoms lead to health care utilization.</td>
<td>Medium Longitudinal sectional log based study. This is appropriate to analyse online user behaviour but less appropriate to determine behaviour the Web.</td>
<td>High The authors describe which data was available, which logs were included/excluded, how they identified logs relating to common symptoms, and how they defined logs showing healthcare utilization intent.</td>
<td>Medium The authors draw conclusions about website and user characteristics that are related to healthcare utilization intent, and demonstrate this using logistic regression analysis. But the measure of healthcare utilization intent may be flawed; it is not clear it searching for healthcare institutions reflects healthcare utilization intent.</td>
<td>Medium The authors use logistic regression to predict healthcare utilization intent queries based on traits of Web pages viewed as well as user characteristics. Beta coefficients were not reported.</td>
<td>Log based study – assumes that certain queries relate to certain topics based on certain algorithms. Some queries may therefore be wrongly classified or the algorithms may not assess what they assume to assess. Specifically: querying for healthcare facilities may not mean intent to visit them, and searching for symptoms may not mean self-diagnosis.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------</td>
</tr>
<tr>
<td>White &amp; Horvitz, 2012</td>
<td>High The aim is described in a clear sentence. The aim was to explore how users search for medical concerns and particularly how these concerns impact on future behaviour, e.g. how this influences focus and attention of future searches.</td>
<td>High Longitudinal log-based study. This is appropriate to study user behaviour over time.</td>
<td>Medium The authors describe how they identified logs related to medical concerns and how they labeled these, and also define clearly the different transitions types examined (e.g. transition from symptom to serious illnesses, or symptoms to benign explanations). Exact number of users/logs not stated.</td>
<td>Medium The authors back up their claims about user behaviour by describing the logs which they previously assigned to categories. However the authors make assumptions that are potentially flawed – e.g. that the user is searching pre-diagnosis, when they may be researching a diagnosed condition.</td>
<td>High Descriptive statistics are used to describe user behaviour (e.g. percentage of logs symptoms related), and logistic regression was used to predict whether a Web page was medical or not.</td>
<td>Log based study – assumes that certain queries relate to certain topics based on certain algorithms. Some queries may therefore be wrongly classified or the algorithms may not assess what they assume to assess.</td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>White &amp; Horvitz, 2013</td>
<td>High</td>
<td>Longitudinal log-based study. This method is suitable to examine result pages of search engines and user behaviour (hovering and clicks).</td>
<td>High</td>
<td>The authors provide descriptive statistics (percentages) to describe results returned by search engines, and significance values to establish whether behaviours of users differ depending on whether snippets relate to serious or benign causes of symptoms.</td>
<td>Medium</td>
<td>The authors use t tests, ANOVA and Tukey-Kramer tests; these are appropriate to establish whether user behaviours differ depending on snippet content. The authors describe in detail how they developed their click prediction model based on the Dynamic Bayesian Network model. Because methods are described alongside results rather than all under a specific heading (‘statistical analyses’) it is difficult to establish exactly which tests the authors used.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Log based study – assumes that certain queries relate to certain topics based on certain algorithms. Some queries may therefore be wrongly classified or the algorithms may not assess what they assume to assess.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Internal validity</td>
<td></td>
</tr>
<tr>
<td>Authors, date</td>
<td>Are the aims and objectives of the research clearly stated?</td>
<td>Is the research design clearly specified and appropriate for the aims and objectives of the research?</td>
<td>Do the researchers provide a clear account of the process by which their findings were reproduced? (methods)</td>
<td>Do the researchers display enough data to support their interpretations and conclusions?</td>
<td>Is the method of analysis appropriate and adequately explicated?</td>
<td>Main limitations</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Yastik, 2017</td>
<td>High The overall aims as well as specific objectives are described clearly.</td>
<td>Low It is clearly specified and appropriate for some of the stated objectives (describing information gathering and user characteristics); however a cross-section survey is not appropriate for the stated objective of determining whether internet usage influences seeking medical care. The appropriate control group for this aim is also missing.</td>
<td>High The authors describe how patients were approached, how the survey was administered, which questions the survey included, and how data were analysed</td>
<td>High Conclusions are based on descriptive data and not overstated.</td>
<td>Medium Yes, for the descriptive elements. However, the model testing the influence of Internet use on ED visits should have been reported in more detail, and the sample size was too small for multiple regression (especially given the number of predictors entered in the model)</td>
<td>Small sample size (N=74), especially considering the target for the regression model was 242. Retrospective self-reported accounts – may be biased and/or inaccurate. Appropriate control group missing as only patients who engaged in Web use were included</td>
</tr>
<tr>
<td>Ybarra &amp; Suman, 2006</td>
<td>High The aim is described in a clear succinct sentence: The aim was to examine which user characteristics predict whether a Web user is likely to contact a health professional.</td>
<td>High A national, cross-sectional, randomly sampled telephone survey. This method is appropriate to establish characteristics of individuals engaging in certain behaviour.</td>
<td>High The authors describe how the survey was conducted, how participants were contacted and recruited, and which questions were asked.</td>
<td>High The authors back up their claims by reporting significance values and odds ratios, percentages of individuals endorsing certain behaviours/characteristics. They do not draw causal conclusions.</td>
<td>High Chi square, t-tests, ANOVA. These tests are appropriate to identify differences between those reporting contact with health professionals and those who did not.</td>
<td>Retrospective self-reported accounts – may be biased and/or inaccurate. The authors do not consistently distinguish between people seeking health information online to diagnose symptoms, and those seeking health information to look up an already received diagnosis.</td>
</tr>
</tbody>
</table>
Appendix 4: NVivo matrix example for the systematic review

Systematic Review (NVivo 11).rvg - NVivo Pro

Younger adults tended to look up more links, leave a web page quickly and be more within-category links. Older adults tended to look up only a subset of links longer time to click a link and leave a web page, and browse more between category links. This is consistent with the notion that younger people used a more bottom and strategy and followed the structure of the interface, clicking on consecutive links within the same category, while older adults used their background knowledge which links to click on, and thus clicked more between categories rather than following the structure of the interface.

Symptoms are tweeted more than searched except potentially embarrassing/stigmatised/sensitive symptoms (e.g. vaginal bleeding, pelvic pain). Averaged together with the vs. decision-maker as a result of they may be the vs. they interact with.

Previous research tended to advise for information to be sensitive to profiles of older adults as more open-ended goals, older adults and information from clinicians that are not.
Appendix 5: NHS Ethics approval letter

Health Research Authority
NRES Committee South West - Central Bristol
Whitefriars
Level 3, Block B
Levin's Mead
Bristol BS1 2NT
Email: nrescommittee.southwest-bristol@nhs.net
Telephone: 01179221380
Fax: 011792220445

20 May 2014

Julia Mueller
PhD Student
University of Manchester
University of Manchester
School of Nursing, Midwifery and Social Work, R.3.330
Oxford Road
M13 9PL

Dear Julia,

Study title: The role of web based information on help seeking of those worried about lung cancer prior to diagnosis
REC reference: 14/SW/0128
IRAS project ID: 143412

The Proportionate Review Sub-committee of the NRES Committee South West - Central Bristol reviewed the above application on 19 May 2014.

We plan to publish your research summary wording for the above study on the NRES website, together with your contact details, unless you expressly withhold permission to do so. Publication will be no earlier than three months from the date of this favourable opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to make a request to postpone publication, please contact the REC Manager Mrs Naaz Nathoo, nrescommittee.southwest-bristol.nhs.net

Ethical opinion

On behalf of the Committee, the sub-committee gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

A Research Ethics Committee established by the Health Research Authority
Participant Information Sheet (PIS)

THE ROLE OF INFORMATION SOURCES BEFORE A DIAGNOSIS OF LUNG CANCER

Julia Mueller, PhD Candidate at the University of Manchester

You are being invited to take part in a research study as part of a PhD degree at the University of Manchester. Please take time to read the following information carefully, ask me any questions you like and take time to decide whether you wish to take part.

What is the purpose of the study?

In the past decade or so, it has become more common for people to use the internet and other information sources to look up their symptoms when they feel unwell, to help them find out if they have a serious condition that needs medical attention. The overall aim of this project is to find out how people with lung cancer use different information sources to get informed about their symptoms before they go see their doctor. We are also interested in why people do not use information sources. The information gathered in this research will be used to help inform the public about lung cancer, and help people decide when they should see their doctor.

Why have I been invited to take part in this study?

We are inviting patients from local cancer centres who were diagnosed with lung cancer in the previous six months.

Do I have to take part?

No, you do not have to take part in this study if you do not want to. It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. Your decision to participate in this study will not be connected to the treatment you are receiving now or in the future. If you decide to take part you are still free to withdraw at any time without giving a reason and without any consequences to your current or future treatment.

What will taking part involve?
**Questionnaire:** Should you decide to participate, you will be asked to complete a questionnaire which will take approximately 10-15 minutes. The questionnaire includes questions about your experiences with your illness before you were diagnosed. There are also questions about any information sources you may or may not have used to get informed about the symptoms you were experiencing. You can fill in the questionnaire now (in the clinic) or you can fill it in at home and then mail it to the research team using the stamped addressed return envelope provided. You can fill the questionnaire in on your own or, if you prefer, you can fill in the questionnaire together with me (the researcher). I would also like to look at relevant sections of your clinical records, for example to find out the exact date when you were first diagnosed.

**Interview (optional):** If you are happy to participate in a further face-to-face interview, I may contact you in a few weeks to arrange an appointment. This part of the study is optional; you can participate in the main study without participating in the interviews. In this interview, we would like to hear some more about how you experienced the time before your diagnosis, and what helped you to form the decision to see your doctor. We are especially interested in hearing your personal views, opinions and experiences. The interview will take approximately one hour and will be audio-recorded. Interviews will take place at a location of your choice where you feel comfortable, for example in your home. If you prefer to have the interview at another location such as your clinic, we can reimburse your travel expenses.

**What are the possible disadvantages and risks of taking part?**

It is not our aim to ask any questions that might be upsetting or stressful to you. However you might be uncomfortable thinking or talking about your illness. If there are any questions you do not wish to answer or you feel you want to withdraw from the study at any point, that is absolutely fine.

**What are the possible benefits of taking part?**

We cannot promise the research will help you but the information we get might help inform other people with lung cancer and help them to seek timely medical attention. Some people find taking part in interviews about their illness helpful.

**How is confidentiality maintained?**
All data will be stored securely at the University of Manchester. Any information collected about you during this study will be kept strictly confidential. Data from the questionnaires and interviews will be made anonymous, so your name will not be included. Documents with your personal information (name and address) will be stored in a locked cabinet in a locked room which only the research team can access. The interviews will be audio-recorded and transcribed word-for-word. After transcription the original tapes will be destroyed. Direct quotes may be used in the write-up of the study, but these will be made anonymous. Data on computers will be encrypted so they can only be accessed with the correct password, to ensure only the research team can access the data. Data from the study will be kept for a minimum of 10 years after the date of any publication which is based upon it, to follow recommended good practice guidelines for research.

Study data and material may be looked at by individuals from the University of Manchester, from regulatory authorities or from the NHS trust, for monitoring and auditing purposes, and this may well include access to personal information.

What will happen if I do not want to carry on with the study?

You can withdraw from the study completely at any time without giving a reason and without any consequences to your current or future treatment. No further data will be collected from the moment you withdraw.

What if there is a problem?

It is unlikely that anything would go wrong. But if there is a problem, you may contact me in the first instance or you can contact my supervisor Professor Chris Todd.

If there are any issues regarding this research that you would prefer not to discuss with me or my supervisor, please contact the Research Practice and Governance Co-ordinator by either writing to ‘The Research Practice and Governance Co-ordinator, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester M13 9PL’, by emailing: research.complaints@manchester.ac.uk, or by telephoning 0161 275 7583 or 275 8093.

What will happen to the results of the study?

The information obtained will be analysed and written up as part of a PhD thesis. The data will also be summarised in the form of research articles which
will be published in a medical journal. Should you like a summary of the results please provide contact details on the questionnaire and we will mail this to you.

Who is funding the research?

This research is funded by the Medical Research Council.

Who has reviewed the study?

All research which involves NHS patients has to be reviewed by the National Health Service Research Ethics Committee (REC). Ethical approval for this study was granted by [name of REC committee granting approval will be inserted here].

Whom can I contact for further information?

If you have any questions or require any additional information about this research project please do not hesitate to contact me or my supervisor at:

**Julia Mueller (chief investigator)**  
*School of Computer Science,  
LF1 Kilburn Building,  
Oxford Road, MANCHESTER,  
M13 9PL, UK*  
*julia.mueller@postgrad.manchester.ac.uk*  
*Tel.: 07449 64 47 17*

**Professor Chris Todd (supervisor)**  
*Professor of Primary Care and Community Health*  
*The School of Nursing, Midwifery and Social Work*  
*University of Manchester*  
*Room 6.314b, Jean McFarlane Building*  
*Oxford Road*  
*Manchester M13 9PL*  
*chris.todd@manchester.ac.uk*  
*Tel.: 0161 306 7865*

If you would like some additional information on lung cancer and available support services (e.g. practical, medical, financial or psychological support), you can contact:

**Macmillan Cancer support**  
*www.macmillan.org.uk*

**The Roy Castle Lung Cancer Foundation**  
*www.roycastle.org*
Macmillan Cancer Support
89 Albert Embankment, London, SE1 7UQ
Phone: 0808 808 00 00

The Roy Castle Centre (Head Office)
Enterprise Way, Wavertree Technology Park
Liverpool, Merseyside, L13 1FB
Phone: 0333 323 7200
Email: foundation@roycastle.org

Should you feel the need to talk to someone about your illness please contact the lung cancer nurse referred to you at diagnosis or to another healthcare practitioner from your treatment team, who can point you to further support services in your clinic.
Appendix 7: Questionnaire/survey

STUDY: THE ROLE OF INFORMATION SOURCES BEFORE A DIAGNOSIS OF LUNG CANCER

Questionnaire

Please take a few minutes to fill out this brief questionnaire about your experiences with your illness before you received your diagnosis of lung cancer.

There are no right or wrong answers to these questions. Just reply according to your own experiences, thoughts or opinions.

Please start by stating some information about yourself:

What is your date of birth? __ __ / __ __ / __ __ __ __  
(D D) / (M M) / (Y Y Y Y )

What is your sex?  
[ ] Male  [ ] Female

Current occupation: ____________________________

SECTION A: Illness experience before the diagnosis.

To start, we would like to ask you a few questions about the period before you received your diagnosis of lung cancer.

A1. Which physical changes or symptoms did you notice before you first went to see the doctor? You can tick as many as you like.

[ ] A persistent / nagging cough
[ ] A long standing cough that got worse
[ ] Coughing up blood
[ ] Breathlessness / Not being able to catch your breath
[ ] Fatigue / constant tiredness / lack of energy
[ ] Weight loss / loss of appetite
[ ] I did not notice any symptoms before I went to see the doctor, the illness was discovered when I went to the doctor for another reason.

[ ] Other, please specify:
A2. How long before you first contacted the doctor did you start noticing symptoms? Please tick one. If you are unsure, just give your best estimate.

- [ ] About one week
- [ ] About two weeks
- [ ] About three weeks
- [ ] About one month
- [ ] About two months
- [ ] About three months
- [ ] About four months
- [ ] About five months
- [ ] About six months
- [ ] More than six months
A3. What thoughts did you have about these symptoms before you sought medical help? You can tick as many as you like, or use the provided space to describe other thoughts you had.

- I did not think these changes meant that I was ill at first.
- I thought these changes were related to another condition (such as a bad cold, asthma, smokers’ cough …)
- I thought that I was ill, but it didn’t seem like it needed professional medical attention at first.
- I thought that I could self-treat using home remedies or over-the-counter medication.
- I was afraid that it might be something serious.
- I felt uncomfortable about seeing a doctor because I have not been leading a very healthy lifestyle.
- I did not want to bother the doctors because I know they are very busy.
- I wanted to make an appointment with my doctor but they did not have an appointment available immediately.
- I wanted to make an appointment with my doctor, but other pressing matters in my life were making this difficult.
- None of the above.
- Other:
### SECTION B: Use of information resources before the diagnosis.

**B1. Did you talk to anyone or use any information sources to help you understand symptoms before you were diagnosed?** Please tick all that apply from the list below:

- [ ] Internet/world-wide Web
- [ ] Talking to family, friends, work colleagues or other people
- [ ] Leaflets
- [ ] Magazines
- [ ] Books
- [ ] Television
- [ ] Radio
- [ ] Medical journals or libraries
- [ ] Public library
- [ ] Other, please specify:
  - [ ] None
B2. Did a friend or a relative talk to anyone or use any information sources to help you understand your symptoms before you were diagnosed? Please tick all that apply from the list below:

- [ ] Internet/world-wide Web
- [ ] Talking to family, friends, work colleagues or other people
- [ ] Leaflets
- [ ] Magazines
- [ ] Books
- [ ] Television
- [ ] Radio
- [ ] Medical journals or libraries
- [ ] Public library
- [ ] Other, please specify:
- [ ] None
B3. If you, or a friend or relative used any information sources to help you understand your symptoms before you were diagnosed, please rate how helpful you found the information you got.

<table>
<thead>
<tr>
<th>Information Source</th>
<th>Very unhelpful</th>
<th>Somewhat unhelpful</th>
<th>Somewhat helpful</th>
<th>Very helpful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet/world-wide Web</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Talking to family, friends, work colleagues or other people</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leaflets</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Magazines</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Books</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Television</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical journals or libraries</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public library</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other, please specify:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other, please specify:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B4. Do you ever use any of the following? You can tick as many as you like.

- [ ] Computer
- [ ] Internet connection at home
- [ ] A mobile phone that can go on the internet (e.g. an Android phone or an iphone)
- [ ] A tablet or an ipad
- [ ] None of the above
B5. Do you own any of the following? You can tick as many as you like.

☐ Computer

☐ A mobile phone that can go on the internet (e.g. an Android phone or an iphone)

☐ A tablet or an ipad

☐ None of the above

B6. How much do you use the internet for any activity during a typical week? If you are unsure, just give an estimate.

____________ hours per week

☐ I never use the internet.

B7. Did you, or a friend or relative, use the internet to look up any of your symptoms before you received your diagnosis of lung cancer?

☐ Yes ☐ No

If you answered ‘No’, please skip to Section D.
Section C: Internet use before the diagnosis

This section is concerned with how you used the internet to research your symptoms. If you did not use the internet, please skip to Section D.

C1. Who searched the internet to look up your symptoms? You can tick as many as you like. If you searched the internet together with someone, please tick “Me” and the other person you searched with.

☐ Me

☐ Husband / wife / partner

☐ Child

☐ Grandchild

☐ Brother / Sister

☐ Other relative

☐ Friend

☐ Work colleague

☐ Other, please specify:
C2. How did you find the information on the internet? You can tick as many as you like.

☐ Google
☐ Bing
☐ Yahoo! Search
☐ Ask
☐ Aol Search
☐ I didn't use a search engine.
☐ Don’t know / Can’t remember.
☐ Other, please specify:

C3. What words did you type to find information about your symptoms? Please list as many as you can remember:

____________________________________________________________________

C4. What type(s) of website(s) did you visit when you were looking up your symptoms? You can tick as many as you like.

☐ NHS websites
☐ WebMD
☐ Yahoo! Health
☐ Patient.co.uk
☐ Netdoctor.co.uk
☐ Discussion forums
☐ Blogs
☐ Don’t know / Can’t remember
☐ Other, please specify:
SECTION D: Demographic information.

Finally, please supply some brief information about yourself.

D1. Would you describe yourself as

☐ Single

☐ Married / In a long-term partnership

☐ Divorced / separated

☐ Widowed

☐ Prefer not to say

D2. What is your highest vocational / educational qualification? (please tick one)

☐ CSE / GCSE / Foundation Diploma / Higher Diploma or equivalent

☐ A levels / AS Level/Highers / Advanced Highers or equivalent

☐ NVQ (any level) or other vocational / work-related qualification

☐ Professional qualification (Example teaching, nursing, accountancy)

☐ Undergraduate degree (Example BA, BSc)

☐ Postgraduate degree (Example MA, MSc, PhD)

☐ Not sure

☐ I don’t have a formal education

☐ Other, please specify:

☐ Prefer not to say
D3. Would you describe yourself as

☐ White

☐ Mixed / multiple ethnic groups

☐ Asian / Asian British

☐ Black / African / Caribbean / Black British

☐ Other ethnic group

☐ Prefer not to say

Thank you for completing this questionnaire.

Please use the space below to make any comments you would like to about the questionnaire.

If you decided to fill in the questionnaire in the clinic, please return it now to the researcher. If you filled the questionnaire in at home, please return it by mail in the stamped addressed envelope provided.

If you would like to talk to a member of the research team about your questionnaire, please feel free to telephone us on 07449 644717.

Should you feel the need to talk to someone about your illness please contact the lung cancer nurse referred to you at diagnosis, who can also point you to further support services in your clinic.

If you indicated that you would be happy to participate in a further interview, we may contact you in a few weeks’ time about this.

Comments:
Appendix 8: Interview protocol

STUDY TITLE: THE ROLE OF INFORMATION SOURCES BEFORE A DIAGNOSIS OF LUNG CANCER

Interview protocol

TO BE FILLED IN BY THE INTERVIEWER PRIOR TO THE INTERVIEW:

☐ The interviewee used the internet: Sections A, B
☐ The interviewee did not use the internet: Sections A, C

THE INTERVIEWEE INDICATED THE USE OF THE FOLLOWING HEALTH INFORMATION SOURCES:

____________________________________________________________________
____________________________________________________________________
____________________________________________________________________
____________________________________________________________________

Introduction: Thank you for taking the time to talk to me today and helping me in my research. You already filled in a questionnaire as part of my study, and I would now like to hear a little more about your personal views, opinions and experiences. Before we start, do you have any questions?

In this interview, I would like to talk about the time before you received your diagnosis of lung cancer. There are no right or wrong answers; I am only interested in hearing about your personal views and opinions.
SECTION A: General questions for all interviewees

A1. I am interested in hearing about the time period before you first went to see a doctor about your symptoms. Can you tell me a little more about how you first noticed that you were ill?

Prompts:

a. When did you first start noticing symptoms?
b. What symptoms did you notice?
c. What made you think you were ill?
d. What did you do about the symptoms at first?
e. When did you decide to seek professional help?

A2. I am interested in understanding what motivated you to seek medical help. Can you tell me a little more about what helped you decide to see a doctor?

Prompts:

a. Was it one of the symptoms that prompted you to see your GP? If yes, which one and why?
b. Did any of the symptoms change? How, and why was it alarming?
c. Did a friend/relative encourage you to see your doctor? Who and how?
d. Did you hear/read some new information that made you think you should see your doctor? If yes, what information and how/why did that prompt you?

A3. Was there anything that stopped you from seeing your doctor at first?

Prompts:

a. This could be something within yourself, such as thoughts or feelings, or an event in your life which prevented you from contacting your doctor.

A4. Did you know anything about lung cancer before you were diagnosed?

Prompts:

a. What did you know?
b. Where did this information come from?
c. Did this have any impact on your decision to seek help?
d. Are you a current or ex-smoker? If ex-smoker, what made you decide to quit?

SECTION B: Questions about internet use

B1. In the questionnaire you completed a few weeks ago you indicated that you used the internet to look up your symptoms before you were diagnosed. How did you go about your search?

Prompts:

a. Did you use a search engine? Which one?

b. What search terms did you use?

c. What type of websites did you use (NHS, charity websites, forum discussions…)?

d. Why did you use these websites?

e. How many websites did you look at?

f. How did you decide which information was trustworthy?

g. How long did you spend searching?

h. Did you discuss the information you found with someone else? With whom and why?

i. Did you have any initial ideas what might be the cause of the symptom?

j. What was the goal of your search?

k. Have you ever been presented with any information about lung cancer? If yes, how / where?

l. Did any health professional ever offer any advice on smoking cessation and risk of lung cancer? If yes, can you tell me a little more about it?

B2. What made you decide to use the internet?

Prompts:

a. How do you generally feel about health information on the internet?

B3. Did you feel like the internet helped you to make sense of your symptoms?

IF YES: Can you tell me more about how this helped you?

IF NO: Can you tell me more about why you did not find this helpful?
B4. What happened after you searched for information on your symptoms?

Prompts:
   a. What did you do next?
   b. Did you search for further information? How and why?
   c. Did you discuss the information you found with others? With whom and why?
   d. How did you feel after the internet search (reassured, scared, confused…)?
   e. How did you feel about your symptoms after you searched the internet?

B5. What effect do you think the information you found had on your decision to contact your doctor?

Prompts:
   a. Did it encourage you to try some self-treatments at home?
   b. Did it encourage you to make an appointment?

B6. How trustworthy do you think the information you found was?

B7. Was there any information you found especially useful?

B8. Was there any information you found especially unhelpful?

B9. How do you think websites could be improved to help people with lung cancer make sense of their symptoms?

   a. Is there any information that was missing that you think would have helped you?
   b. Is there anything specific you think should be emphasized?
SECTION C: Questions about not using the internet

C1. My study is about researching how people use the internet to research their symptoms before they contact their doctor. In your questionnaire, you indicated that you did not use the internet to make sense of your symptoms before your diagnosis. I am interested in hearing why this is. Can you tell me why?

   a. Did you prefer to use other sources instead? If yes, what other sources and what advantages do they have over the internet?
   b. Did you prefer to talk to others? If yes, with whom and why?
   c. How do you generally feel about health information on the internet?
   d. Do you use the internet for other purposes?

SECTION D: Questions about other health information sources.

E1. In the questionnaire you completed a few weeks ago, you indicated that you used ......................... and ......................... to help you understand your symptoms. What made you decide to use ......................... / .........................?

   Prompts:

   a. Do you generally trust information from this source? Why?

E2. Did you feel like the information you found helped you to make sense of your symptoms?

   IF YES: Can you tell me more about how this helped you?

   IF NO: Can you tell me more about why you did not find this helpful?

E3. What happened after you searched for information on your symptoms?

   Prompts:

   f. What did you do next?
   g. Did you search for further information? How and why?
   h. Did you discuss the information you found with others? With whom and why?
i. How did you feel about your symptoms after you searched for information using this source?

E4. What effect do you think the information you found had on your decision to contact your doctor?

Prompts:

c. Did it encourage you to try some self-treatments at home?

d. Did it encourage you to make an appointment?

E5. How trustworthy do you think the information you found was?

Thank you for participating in this research project and taking the time to talk to me today. Do you have any more questions you would like to ask me?
Appendix 9: NVivo matrix example for Framework Analysis

![NVivo Matrix Example](image-url)
Appendix 10: UREC approval letter

Ref: ethics/15353

Ms Julia Mueller
PhD Student
School of Nursing, Midwifery and Social Work
Faculty of Medical and Human Sciences
University of Manchester
M13 9PL
Julia.mueller@manchester.ac.uk

30 September 2015

Dear Ms Mueller

RE: Ref 15353: Evaluation of a Web-based intervention to encourage early help-seeking in people with symptoms associated with lung cancer: A feasibility study

Research Ethics Committee 1

The University Research Ethics Committee 1 met on 10th September 2015 to discuss the above study (at which your supervisor Dr Catherine Jay was in attendance). I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form and supporting documentation, as submitted to and approved by the Committee.

This approval is effective for a period of five years. If the project continues beyond that period an application for amendment must be submitted for review. Likewise, any proposed changes to the way the research is conducted must be approved via the amendment process (see below). Failure to do so could invalidate the insurance and constitute research misconduct.

You are reminded that, in accordance with University policy, any data carrying personal identifiers must be encrypted when not held on a secure university computer or kept securely as a hard copy in a location which is accessible only to those involved with the research.

Reporting Requirements:

You are required to report to us the following:

1. Amendments
2. Breaches and adverse events
3. Notification of progress/End of the Study

Feedback
EVALUATION OF A TOOL TO HELP PEOPLE APPRAISE SYMPTOMS AND DECIDE WHEN TO SEEK MEDICAL HELP

Julia Mueller, PhD Candidate at the University of Manchester

You are being invited to take part in a research study as part of a PhD degree at the University of Manchester. Please take time to read the following information carefully and take time to decide whether you wish to take part.

What is the purpose of the study?
For certain symptoms that may be associated with more serious illnesses, such as lung cancer, it is especially important to provide clear and helpful information to help people decide when they need to see a doctor about their symptoms. We have therefore developed a tool to help people appraise their symptoms and to encourage people to see their GP if necessary. The aim of this study is to find out how people use this tool and whether it is effective. This tool is not designed to help you diagnose yourself. It will help you decide whether you need to see a GP about your symptoms. If this tool does suggest that you should see a GP, this does not necessarily mean that you have lung cancer. It just means it would be better to get your symptoms checked out.

Why have I been invited to take part in this study?
We are inviting anyone to participate who is above the age of 18, lives in the UK and is experiencing any of the symptoms listed below. We are also inviting those above the age of 18 and living in the UK who know anyone who has been experiencing any of the following symptoms, for the past few weeks:

- A cough (whether dry and tickly or productive with phlegm)
- A long-standing cough that changes or gets worse
- Feeling out of breath
- Discomfort in the chest, shoulder or back
- Coughing up phlegm with a blood in it (even if just a few specks) / spitting blood
- Changes in your voice
- Unexplained weight loss or unexplained loss of appetite
- Swelling of face and neck
- Persistent / recurring chest infection
- Tiredness or lack of energy
- Any changes in the appearance of the fingers or fingernails (such as a softening of the nailbed, stronger than normal curving of the fingernails, or thickening of the fingertips so that the shape looks like an upside down spoon)

Do I have to take part?

No, you do not have to take part in this study if you do not want to. It is up to you to decide whether or not to take part. If you decide to take part you are still free to withdraw at any time without giving a reason.

What will taking part involve?

The study will take place online. If you decide to take part, you will first be asked a few basic questions about yourself, or about the person who has been experiencing the symptoms. You will then be guided through the tool, which will provide information on symptoms and risk factors. In the end, you will receive advice on whether you should make an appointment to see your GP. Finally, you will be asked to fill in a questionnaire. Taking part in this study will take approximately 15-20 minutes. You will not be able to save the questionnaire and return to it at a later time so you will need to complete it in one go.

What are the possible disadvantages and risks of taking part?

It is not our aim to ask any questions or provide any information that might be upsetting or stressful to you. If there are any questions you do not wish to answer or you feel you want to withdraw from the study at any point, that is absolutely fine. If you feel distressed or worried after taking part in this study, it will most likely be best if you contact your GP and speak to your GP about your concerns. If you would like further support or advice, you can also call MacMillan Cancer support for free on 0808 808 00 00.

What are the possible benefits of taking part?

We cannot promise the research will help you but you may find the information provided here useful in deciding whether you need to see a GP about your symptoms, or whether you should encourage someone else to see a GP.

How is confidentiality maintained?

All data will be stored securely at the University of Manchester. Any information collected about you during this study will be kept strictly confidential. Data from the questionnaires and interviews will be made anonymous, so your name will not be included. All data will be stored in an encrypted format so they can only be accessed with the correct password, to ensure only the research team can access the data. Data
from the study will be kept for a minimum of 10 years after the date of any publication which is based upon it, to follow recommended good practice guidelines for research.

Study data and material may be looked at by individuals from the University of Manchester or from regulatory authorities, for monitoring and auditing purposes, and this may well include access to personal information.

What will happen if I do not want to carry on with the study?

You can withdraw from the study at any time without giving a reason. From the moment you withdraw, we will not collect any further data from you.

What if I want to complain?

If there is a problem, you may contact me (Julia) in the first instance or you can contact my supervisor Professor Chris Todd.

If there are any issues regarding this research that you would prefer not to discuss with me or my supervisor, please contact the Research Practice and Governance Co-ordinator by either writing to ‘The Research Practice and Governance Co-ordinator, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester M13 9PL’, by emailing: research.complaints@manchester.ac.uk, or by telephoning 0161 275 7583 or 275 8093.

What will happen to the results of the study?

The data obtained will be analysed and written up as part of a PhD thesis. The data will also be summarised in the form of research articles which will be published in a scientific journal.

Who is funding the research?

This research is funded by the Medical Research Council.

Who has reviewed the study?

Ethical approval for this study was granted by The University of Manchester Research Ethics committee (Ref #15353).

Whom can I contact for further information?

If you have any questions or require any additional information about this research project please do not hesitate to contact me or my supervisor at:

Julia Mueller (chief investigator)  Professor Chris Todd (supervisor)
Professor of Primary Care and Community Health

Manchester M13 9PL', by emailing: research.complaints@manchester.ac.uk, or by telephoning 0161 275 7583 or 275 8093.
If you would like some additional information on lung cancer you can contact:

**Macmillan Cancer support**

www.macmillan.org.uk

Macmillan Cancer Support
89 Albert Embankment, London, SE1 7UQ
Phone: 0808 808 00 00

**The Roy Castle Lung Cancer Foundation**

www.roycastle.org

The Roy Castle Centre (Head Office)
Enterprise Way, Wavertree Technology Park
Liverpool, Merseyside, L13 1FB
Phone: 0333 323 7200
Email: foundation@roycastle.org
Appendix 12: Initial survey (Feasibility study)

Do you live in the UK?
- Yes
- No

What is your age?
- Under 40
- 40 and over

What is your gender?
- Male
- Female

Are you a smoker?
- Yes, I currently smoke
- No, but I used to smoke
- No, I have never smoked

Over the past few weeks, have you been coughing?
- Yes, a dry tickly cough
- Yes, a cough with phlegm
- Yes, but I have always had a cough
- No, I have not been coughing at all

Have you noticed any changes in a long-standing cough you've had?
- Yes, the cough has changed or gotten worse
- No, no changes in a cough
In the past few weeks, have you coughed up any phlegm with blood in it, or spit blood (even if just a few specks)?

- Yes
- No

In the past few weeks, have you felt more breathless than usual?

- Yes
- No

In the past few weeks, have you had any discomfort in your chest, shoulders or back?

- Yes
- No

In the past few weeks, have you noticed any changes in your voice (such as hoarseness, sounding deeper, weaker or quivering)?

- Yes
- No

In the past few weeks, have you noticed any noises when breathing (such as wheezing or harsh breath sounds)?

- Yes
- No

In the past few weeks, have you unintentionally (without dieting) lost any weight?

- Yes
- No

In the past few weeks, have you felt more tired than usual?

- Yes
- No
In the past few weeks, have you noticed any swellings or lumps around the face or neck area?

☐ Yes
☐ No

In the past months, have you had recurring or persistent chest infections?

☐ Yes
☐ No

In the past few weeks, have you noticed any changes in the appearance of your fingers or fingernails?

☐ Yes
☐ No

Have you already been to a doctor about your symptoms?

☐ Yes
☐ No

Note: The questionnaire displayed additional questions for some symptoms if they were answered with ‘yes’, example shown below. For symptoms like haemoptysis, where duration does not affect the advice of whether help should be sought, no additional questions were asked.

Over the past few weeks, have you been coughing?

☐ Yes, a dry tickly cough
☐ Yes, a cough with phlegm
☐ Yes, but I have always had a cough
☐ No, I have not been coughing at all

For how long have you been coughing?

☐ Less than 3 weeks
☐ 3 weeks or longer

Is the cough very severe (strong / intense)?

☐ Yes
☐ No
Appendix 13: TPB-components item

Which of the following statements do you agree with most?

- “Making an appointment to see a doctor about these symptoms would be pointless.”
- “It is important whether others (family and friends) think it is necessary to see a doctor about these symptoms.”
- “It is important whether doctors think the symptoms are worth investigating.”
- “Getting an appointment to report these symptoms to a doctor would be difficult.”
Appendix 14: Button-click measure (Feasibility study)

Button-click measure

Would you like to find out how to book an appointment with your doctor now?

Yes » No »

© Julia Mueller 2015 - E: julia.mueller-at-manchester.ac.uk T: +44 161 275 7821 - Participant information sheet
Appendix 15: Intention measure (Feasibility study)

Intention measure

Do you intend to make an appointment with your doctor to have your symptoms checked?

☐ Yes  ☐ No
Note: The education measure showed further details when users hovered over each option, to help users decide which option to select.
Appendix 17: TPB questionnaire

TPB Questionnaire

I would intend to make an appointment to see a doctor about these symptoms.

For me to make an appointment to see a doctor about these symptoms would be

Making an appointment to see a doctor about these symptoms would waste my doctor’s time.

Wasting my doctor’s time is

Making an appointment to see a doctor about these symptoms is important to check if there is a serious cause.

Checking for a serious cause is
Making an appointment to see a doctor about these symptoms, even if they are unlikely to be caused by lung cancer, can help to set your mind at rest.

Setting your mind at rest is

Making an appointment to see a doctor about these symptoms can help to detect lung cancer early and treat it more effectively.

Detecting lung cancer early and treating it more effectively is

Most people who are important to me would want me to make an appointment to see a doctor about these symptoms.

My family and friends would want me to make an appointment to see a doctor about these symptoms

My doctors would want me to make an appointment to see a doctor about these symptoms
When it comes to matters of health, I want to do what my doctors think I should do.

I expect that my doctor’s practice will not have many appointments available.

My doctor’s practice not having many appointments available will prevent me from making an appointment to see a doctor about these symptoms.

I am confident that I could make an appointment to see a doctor about these symptoms.

For me to make an appointment to see a doctor about these symptoms would be
Appendix 18: Risk perception scale

How likely do you feel it is that you will develop lung cancer in your lifetime?

Very unlikely

1 2 3 4 5 6 7 Very likely
Appendix 19: When2Go evaluation

Would you recommend the use of this website to a friend or family member?

☐ Yes

☐ No
Appendix 20: Initial survey (Trial 1)

Initial survey

Do you live in the UK?

☐ Yes
☐ No

What is your age?

☐ Under 40
☐ 40 and over

What is your gender?

☐ Male
☐ Female

Do you smoke?

☐ Yes, I currently smoke
☐ No, but I used to smoke
☐ No, I have never smoked

Do you have any of the following conditions? (select all that apply)

☐ asthma
☐ chronic obstructive pulmonary disease (COPD)
☐ cystic fibrosis
☐ pulmonary fibrosis
☐ a respiratory infection
☐ sleep apnoea
☐ tuberculosis (TB)
☐ other respiratory/lung conditions
☐ none of the above

Over the past few weeks, have you been coughing?
☐ Yes, a dry tickly cough
☐ Yes, a cough with phlegm
☐ Yes, but I have always had a cough
☐ No, I have not been coughing at all

Have you noticed any changes in a long-standing cough you’ve had?
☐ Yes, the cough has changed or gotten worse
☐ No, no changes in a cough

In the past few weeks, have you coughed up any phlegm with blood in it, or spit blood (even if just a few specks)?
☐ Yes
☐ No

In the past few weeks, have you felt more breathless than usual?
☐ Yes
☐ No

In the past few weeks, have you had any discomfort in your chest, shoulders or back?
☐ Yes
☐ No

In the past few weeks, have you noticed any changes in your voice (such as hoarseness, sounding deeper, weaker or quivering)?
Yes

No

In the past few weeks, have you noticed any noises when breathing (such as wheezing or harsh breath sounds)?

Yes

No

In the past few weeks, have you unintentionally (without dieting) lost any weight?

Yes

No

In the past few weeks, have you felt more tired than usual?

Yes

No

In the past few weeks, have you noticed any swellings or lumps around the face or neck area?

Yes

No

In the past months, have you had recurring or persistent chest infections?

Yes

No

In the past few weeks, have you noticed any changes in the appearance of your fingers or fingernails?

Yes

No

Have you already been to a doctor about your symptoms?
☐ Yes
☐ No
Appendix 21: Intention measure (Trial 1)

“Do you intend to see a doctor about your symptoms?”

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No, definitely not</td>
<td>Yes, definitely</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Proxies:

“Do you intend to encourage your friend/relative to see a doctor about their symptoms?”

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No, definitely not</td>
<td>Yes, definitely</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 22: Next steps measure (Trial 1)

What would you like to do now?

- Make an appointment with my GP
- Register with a GP
- Change to a different GP
- None of these
You are being invited to take part in a research study as part of a PhD degree at the University of Manchester. Please take time to read the following information carefully and take time to decide whether you wish to take part.

What is the purpose of the study?
For certain symptoms that may be associated with more serious illnesses, such as lung cancer, it is especially important to provide clear and helpful information to help people decide when they need to see a doctor about their symptoms. We have therefore developed a tool to help people appraise their symptoms and to encourage people to see their GP if necessary. The aim of this study is to find out how people use this tool and whether it is effective. This tool is not designed to help you diagnose yourself. It will help you decide whether you need to see a GP about your symptoms. If this tool does suggest that you should see a GP, this does not necessarily mean that you have lung cancer. It just means it would be better to get your symptoms checked out.

Why have I been invited to take part in this study?
We are inviting anyone to participate who is above the age of 18, lives in the UK and is experiencing any of the symptoms listed below. We are also inviting those above the age of 18 and living in the UK who know anyone who has been experiencing any of the following symptoms, for the past few weeks:

- A cough (whether dry and tickly or productive with phlegm)
- A long-standing cough that changes or gets worse
- Feeling out of breath
- Discomfort in the chest, shoulder or back
- Coughing up phlegm with a blood in it (even if just a few specks) / spitting blood
- Changes in your voice
- Unexplained weight loss or unexplained loss of appetite
- Swelling of face and neck
- Persistent / recurring chest infection
- Tiredness or lack of energy
- Any changes in the appearance of the fingers or fingernails (such as a softening of the nailbed, stronger than normal curving of the fingernails, or thickening of the fingertips so that the shape looks like an upside down spoon)

**Do I have to take part?**

No, you do not have to take part in this study if you do not want to. It is up to you to decide whether or not to take part. If you decide to take part you are still free to withdraw at any time without giving a reason.

**What will taking part involve?**

The study will take place online. If you decide to take part, you will first be asked a few basic questions about yourself, or about the person who has been experiencing the symptoms. You will then be guided through the tool, which will provide information on symptoms and risk factors. In the end, you will receive advice on whether you should make an appointment to see your GP. Finally, you will be asked to fill in a questionnaire. Taking part in this study will take approximately 15-20 minutes. You will not be able to save the questionnaire and return to it at a later time so you will need to complete it in one go.

**What are the possible disadvantages and risks of taking part?**

It is not our aim to ask any questions or provide any information that might be upsetting or stressful to you. If there are any questions you do not wish to answer or you feel you want to withdraw from the study at any point, that is absolutely fine. If you feel distressed or worried after taking part in this study, it will most likely be best if you contact your GP and speak to your GP about your concerns. If you would like further support or advice, you can also call MacMillan Cancer support for free on 0808 808 00 00.

**What are the possible benefits of taking part?**

We cannot promise the research will help you but you may find the information provided here useful in deciding whether you need to see a GP about your symptoms, or whether you should encourage someone else to see a GP.

**How is confidentiality maintained?**

All data will be stored securely at the University of Manchester. Any information collected about you during this study will be kept strictly confidential. Data from the questionnaires and interviews will be made anonymous, so your name will not be included. All data will be stored in an encrypted format so they can only be accessed with the correct password, to ensure only the research team can access the data. Data
from the study will be kept for a minimum of 10 years after the date of any publication which is based upon it, to follow recommended good practice guidelines for research.

Study data and material may be looked at by individuals from the University of Manchester or from regulatory authorities, for monitoring and auditing purposes, and this may well include access to personal information.

**What will happen if I do not want to carry on with the study?**

You can withdraw from the study at any time without giving a reason. From the moment you withdraw, we will not collect any further data from you.

**What if I want to complain?**

If there is a problem, you may contact me (Julia) in the first instance or you can contact my supervisor Professor Chris Todd.

If there are any issues regarding this research that you would prefer not to discuss with me or my supervisor, please contact the Research Practice and Governance Co-ordinator by either writing to ‘The Research Practice and Governance Co-ordinator, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester M13 9PL’, by emailing: research.complaints@manchester.ac.uk, or by telephoning 0161 275 7583 or 275 8093.

**What will happen to the results of the study?**

The data obtained will be analysed and written up as part of a PhD thesis. The data will also be summarised in the form of research articles which will be published in a scientific journal.

**Who is funding the research?**

This research is funded by the Medical Research Council.

**Who has reviewed the study?**

Ethical approval for this study was granted by The University of Manchester Research Ethics committee (Ref #15353).

**Whom can I contact for further information?**

If you have any questions or require any additional information about this research project please do not hesitate to contact me or my supervisor at:
If you would like some additional information on lung cancer you can contact:

**Macmillan Cancer support**  
www.macmillan.org.uk  
Macmillan Cancer Support  
89 Albert Embankment, London, SE1 7UQ  
Phone: 0808 808 00 00

**The Roy Castle Lung Cancer Foundation**  
www.roycastle.org  
The Roy Castle Centre (Head Office)  
Enterprise Way, Wavertree Technology Park  
Liverpool, Merseyside, L13 1FB  
Phone: 0333 323 7200  
Email: foundation@roycastle.org
Appendix 24: Re-run of non-parametric ANOVAs with randomly selected groups

**Intention**

(N=200)

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>27.38</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>3.84</td>
<td>0.924</td>
</tr>
<tr>
<td>Rows</td>
<td>2450.00</td>
<td>1</td>
<td>0.81</td>
<td>3.84</td>
<td>3.84</td>
<td>0.367</td>
</tr>
<tr>
<td>Interaction</td>
<td>162.00</td>
<td>1</td>
<td>0.05</td>
<td>3.84</td>
<td>3.84</td>
<td>0.817</td>
</tr>
<tr>
<td>Within</td>
<td>596714.62</td>
<td>196</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>199</td>
<td>3011.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>941.78</td>
<td>1</td>
<td>0.31</td>
<td>3.84</td>
<td>3.84</td>
<td>0.576</td>
</tr>
<tr>
<td>Rows</td>
<td>598.58</td>
<td>1</td>
<td>0.20</td>
<td>3.84</td>
<td>3.84</td>
<td>0.656</td>
</tr>
<tr>
<td>Interaction</td>
<td>100.82</td>
<td>1</td>
<td>0.03</td>
<td>3.84</td>
<td>3.84</td>
<td>0.855</td>
</tr>
<tr>
<td>Within</td>
<td>597492.32</td>
<td>196</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>199</td>
<td>3010.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>369.92</td>
<td>1</td>
<td>0.12</td>
<td>3.84</td>
<td>3.84</td>
<td>0.726</td>
</tr>
<tr>
<td>Rows</td>
<td>1250.00</td>
<td>1</td>
<td>0.42</td>
<td>3.84</td>
<td>3.84</td>
<td>0.519</td>
</tr>
<tr>
<td>Interaction</td>
<td>180.50</td>
<td>1</td>
<td>0.06</td>
<td>3.84</td>
<td>3.84</td>
<td>0.807</td>
</tr>
<tr>
<td>Within</td>
<td>597349.58</td>
<td>196</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>199</td>
<td>3010.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>142.80</td>
<td>1</td>
<td>0.05</td>
<td>3.84</td>
<td>3.84</td>
<td>0.828</td>
</tr>
<tr>
<td>Rows</td>
<td>1872.72</td>
<td>1</td>
<td>0.62</td>
<td>3.84</td>
<td>3.84</td>
<td>0.431</td>
</tr>
<tr>
<td>Interaction</td>
<td>244.21</td>
<td>1</td>
<td>0.08</td>
<td>3.84</td>
<td>3.84</td>
<td>0.776</td>
</tr>
<tr>
<td>Within</td>
<td>598654.77</td>
<td>196</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>199</td>
<td>3019.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TPB completers (N=120)

**TPB Intention**

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>1116.30</td>
<td>1</td>
<td>1.06</td>
<td>3.84</td>
<td>0.303</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>95.41</td>
<td>1</td>
<td>0.09</td>
<td>3.84</td>
<td>0.763</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>460.21</td>
<td>1</td>
<td>0.44</td>
<td>3.84</td>
<td>0.508</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>123394.58</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>119</td>
<td></td>
<td>1050.98</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>127.10</td>
<td>1</td>
<td>0.12</td>
<td>3.84</td>
<td>0.727</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>35.75</td>
<td>1</td>
<td>0.03</td>
<td>3.84</td>
<td>0.853</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>308.80</td>
<td>1</td>
<td>0.30</td>
<td>3.84</td>
<td>0.586</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>123416.29</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>119</td>
<td></td>
<td>1041.08</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>97.20</td>
<td>1</td>
<td>0.09</td>
<td>3.84</td>
<td>0.759</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>297.67</td>
<td>1</td>
<td>0.29</td>
<td>3.84</td>
<td>0.591</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>0.21</td>
<td>1</td>
<td>0.00</td>
<td>3.84</td>
<td>0.989</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>122385.92</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>119</td>
<td></td>
<td>1031.77</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>76.80</td>
<td>1</td>
<td>0.07</td>
<td>3.84</td>
<td>0.786</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>12.03</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>0.914</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>403.33</td>
<td>1</td>
<td>0.39</td>
<td>3.84</td>
<td>0.533</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>123187.83</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>119</td>
<td></td>
<td>1039.33</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TPB Behavioural attitude**

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>1407.67</td>
<td>1</td>
<td>1.19</td>
<td>3.84</td>
<td>0.276</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>54.67</td>
<td>1</td>
<td>0.05</td>
<td>3.84</td>
<td>0.830</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>8.53</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>0.932</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>139798.12</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>119</td>
<td></td>
<td>1187.13</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>134.41</td>
<td>1</td>
<td>0.12</td>
<td>3.84</td>
<td>0.727</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>15.41</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>0.906</td>
<td></td>
</tr>
</tbody>
</table>
### Interaction

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>( s^2 )</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>264.03</td>
<td>1</td>
<td>0.24</td>
<td>3.84</td>
<td>0.623</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>232.41</td>
<td>1</td>
<td>0.21</td>
<td>3.84</td>
<td>0.644</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>249.41</td>
<td>1</td>
<td>0.23</td>
<td>3.84</td>
<td>0.632</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>129014.15</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>119</td>
<td>1106.74</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TBP Subjective norm

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>( s^2 )</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>1976.41</td>
<td>1</td>
<td>2.23</td>
<td>3.84</td>
<td>0.135</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>2367.41</td>
<td>1</td>
<td>2.67</td>
<td>3.84</td>
<td>0.102</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>208.03</td>
<td>1</td>
<td>0.24</td>
<td>3.84</td>
<td>0.628</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>100772.15</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>119</td>
<td>885.08</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Source of variation

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>( s^2 )</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>2557.63</td>
<td>1</td>
<td>2.89</td>
<td>3.84</td>
<td>0.089</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>1992.67</td>
<td>1</td>
<td>2.25</td>
<td>3.84</td>
<td>0.133</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>63.08</td>
<td>1</td>
<td>0.07</td>
<td>3.84</td>
<td>0.789</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>100544.12</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>119</td>
<td>883.68</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Source of variation

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>( s^2 )</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>755.01</td>
<td>1</td>
<td>0.87</td>
<td>3.84</td>
<td>0.352</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>2910.67</td>
<td>1</td>
<td>3.34</td>
<td>3.84</td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>235.20</td>
<td>1</td>
<td>0.27</td>
<td>3.84</td>
<td>0.603</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>99744.62</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>119</td>
<td>870.97</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Source of variation

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>( s^2 )</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>2557.63</td>
<td>1</td>
<td>2.89</td>
<td>3.84</td>
<td>0.089</td>
<td></td>
</tr>
</tbody>
</table>

341
TPB Composite: “Making an appointment to see a doctor about these symptoms would waste my doctor’s time. “

Source of variation | SS   | df | s²  | Hcalc | Hcrit | p   
Columns            | 1178.13 | 1   | 1.00 | 3.84  | 0.318 
Rows               | 2201.63 | 1   | 1.86 | 3.84  | 0.172 
Interaction        | 36.30   | 1   | 0.03 | 3.84  | 0.861 
Within             | 137194.93 | 116 |      |       |      
Total              | 119     | 119 | 1181.61 |       |      

Source of variation | SS   | df | s²  | Hcalc | Hcrit | p   
Columns            | 1274.01 | 1   | 1.08 | 3.84  | 0.299 
Rows               | 2236.03 | 1   | 1.89 | 3.84  | 0.169 
Interaction        | 118.01  | 1   | 0.10 | 3.84  | 0.752 
Within             | 137032.45 | 116 |      |       |      
Total              | 119     | 119 | 1182.02 |       |      

Source of variation | SS   | df | s²  | Hcalc | Hcrit | p   
Columns            | 143.01  | 1   | 0.12 | 3.84  | 0.727 
Rows               | 5096.03 | 1   | 4.33 | 3.84  | 0.037 
Interaction        | 238.01  | 1   | 0.20 | 3.84  | 0.653 
Within             | 134567.45 | 116 |      |       |      
Total              | 119     | 119 | 1176.84 |       |      

Source of variation | SS   | df | s²  | Hcalc | Hcrit | p   
Columns            | 1274.01 | 1   | 1.08 | 3.84  | 0.299 
Rows               | 2236.03 | 1   | 1.89 | 3.84  | 0.169 
Interaction        | 118.01  | 1   | 0.10 | 3.84  | 0.752 
Within             | 137032.45 | 116 |      |       |      
Total              | 119     | 119 | 1182.02 |       |      

TPB Composite: “Making an appointment to see a doctor about these symptoms is important to check if there is a serious cause.”

Source of variation | SS   | df | s²  | Hcalc | Hcrit | p   
Columns            | 1725.21 | 1   | 1.47 | 3.84  | 0.225 
Rows               | 5644.41 | 1   | 4.82 | 3.84  | 0.028 
Interaction        | 20.83   | 1   | 0.02 | 3.84  | 0.894 
Within             | 132027.55 | 116 |      |       |      

342
<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1171.58</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1182.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Columns</td>
<td>1274.01</td>
<td>1</td>
<td>1.08</td>
<td>3.84</td>
<td>0.299</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>2236.03</td>
<td>1</td>
<td>1.89</td>
<td>3.84</td>
<td>0.169</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>118.01</td>
<td>1</td>
<td>0.10</td>
<td>3.84</td>
<td>0.752</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>137032.45</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1182.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1160.74</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Columns</td>
<td>440.83</td>
<td>1</td>
<td>0.38</td>
<td>3.84</td>
<td>0.538</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>1888.13</td>
<td>1</td>
<td>1.63</td>
<td>3.84</td>
<td>0.202</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>512.53</td>
<td>1</td>
<td>0.44</td>
<td>3.84</td>
<td>0.506</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>135287.00</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1160.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1174.97</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Columns</td>
<td>790.53</td>
<td>1</td>
<td>0.67</td>
<td>3.84</td>
<td>0.412</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>3488.41</td>
<td>1</td>
<td>2.97</td>
<td>3.84</td>
<td>0.085</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>81.67</td>
<td>1</td>
<td>0.07</td>
<td>3.84</td>
<td>0.792</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>135460.88</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1174.97</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TPB Composite:** “Making an appointment to see a doctor about these symptoms, even if they are unlikely to be caused by lung cancer, can help to set your mind at rest.”

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>969.09</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Columns</td>
<td>496.13</td>
<td>1</td>
<td>0.51</td>
<td>3.84</td>
<td>0.474</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>1267.50</td>
<td>1</td>
<td>1.31</td>
<td>3.84</td>
<td>0.253</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>22.53</td>
<td>1</td>
<td>0.02</td>
<td>3.84</td>
<td>0.879</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>113535.83</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>969.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>936.93</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Columns</td>
<td>6.08</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>0.936</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>598.53</td>
<td>1</td>
<td>0.64</td>
<td>3.84</td>
<td>0.424</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>3.67</td>
<td>1</td>
<td>0.00</td>
<td>3.84</td>
<td>0.950</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>110886.72</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>936.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>648</td>
<td>119</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Columns</td>
<td>192.53</td>
<td>1</td>
<td>0.21</td>
<td>3.84</td>
<td>0.648</td>
<td></td>
</tr>
</tbody>
</table>

343
Rows 165.67 1 0.18 3.84 0.672
Interaction 66.01 1 0.07 3.84 0.789
Within 109728.78 116
Total 119 925.66

Source of variation SS df s² Hcalc Hcrit p
Columns 6.08 1 0.01 3.84 0.936
Rows 598.53 1 0.64 3.84 0.424
Interaction 3.67 1 0.00 3.84 0.950
Within 110886.72 116
Total 119 936.93

TPB Composite: “Making an appointment to see a doctor about these symptoms can help to detect lung cancer early and treat it more effectively.”

Source of variation SS df s² Hcalc Hcrit p
Columns 541.88 1 0.65 3.84 0.421
Rows 891.08 1 1.07 3.84 0.302
Interaction 963.33 1 1.15 3.84 0.283
Within 97085.22 116
Total 119 835.98

Source of variation SS df s² Hcalc Hcrit p
Columns 410.70 1 0.47 3.84 0.493
Rows 452.41 1 0.52 3.84 0.472
Interaction 1159.41 1 1.32 3.84 0.250
Within 102149.98 116
Total 119 875.40

Source of variation SS df s² Hcalc Hcrit p
Columns 853.33 1 1.00 3.84 0.316
Rows 140.83 1 0.17 3.84 0.684
Interaction 598.53 1 0.70 3.84 0.401
Within 99499.30 116
Total 119 849.51

Source of variation SS df s² Hcalc Hcrit p
Columns 410.70 1 0.47 3.84 0.493
Rows 452.41 1 0.52 3.84 0.472
Interaction 1159.41 1 1.32 3.84 0.250
Within 102149.98 116
Total 119 875.40
TPB Composite: “My family and friends would want me to make an appointment to see a doctor about these symptoms”

Source of variation | SS    | df | s2  | Hcalc | Hcrit | p
--- | --- | --- | --- | --- | --- | ---
Columns          | 3030.08 | 1   | 2.57 | 3.84 | 0.109
Rows             | 1976.41 | 1   | 1.68 | 3.84 | 0.195
Interaction      | 1092.03 | 1   | 0.93 | 3.84 | 0.335
Within           | 133937.98 | 116 |     |      |      |      
Total            | 119 | 1176.78

Source of variation | SS    | df | s2  | Hcalc | Hcrit | p
--- | --- | --- | --- | --- | --- | ---
Columns          | 1373.63 | 1   | 1.16 | 3.84 | 0.281
Rows             | 3520.83 | 1   | 2.98 | 3.84 | 0.084
Interaction      | 832.13 | 1   | 0.70 | 3.84 | 0.401
Within           | 134743.40 | 116 |     |      |      |      
Total            | 119 | 1180.42

Source of variation | SS    | df | s2  | Hcalc | Hcrit | p
--- | --- | --- | --- | --- | --- | ---
Columns          | 907.50 | 1   | 0.77 | 3.84 | 0.380
Rows             | 4002.08 | 1   | 3.40 | 3.84 | 0.065
Interaction      | 1407.67 | 1   | 1.20 | 3.84 | 0.274
Within           | 133733.75 | 116 |     |      |      |      
Total            | 119 | 1176.90

Source of variation | SS    | df | s2  | Hcalc | Hcrit | p
--- | --- | --- | --- | --- | --- | ---
Columns          | 1373.63 | 1   | 1.16 | 3.84 | 0.281
Rows             | 3520.83 | 1   | 2.98 | 3.84 | 0.084
Interaction      | 832.13 | 1   | 0.70 | 3.84 | 0.401
Within           | 134743.40 | 116 |     |      |      |      
Total            | 119 | 1180.42

TPB Composite: “My doctors would want me to make an appointment to see a doctor about these symptoms”

Source of variation | SS    | df | s2  | Hcalc | Hcrit | p
--- | --- | --- | --- | --- | --- | ---
Columns          | 1548.01 | 1   | 1.34 | 3.84 | 0.246
Rows             | 124.03 | 1   | 0.11 | 3.84 | 0.743
Interaction      | 1.41 | 1   | 0.00 | 3.84 | 0.972
Within           | 135343.55 | 116 |     |      |      |      
Total            | 119 | 1151.40
### Source of variation

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>869.41</td>
<td>1</td>
<td>0.75</td>
<td>3.84</td>
<td>0.386</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>57.41</td>
<td>1</td>
<td>0.05</td>
<td>3.84</td>
<td>0.824</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>120.00</td>
<td>1</td>
<td>0.10</td>
<td>3.84</td>
<td>0.748</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>136758.18</td>
<td>116</td>
<td>1158.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>119</td>
<td>1158.03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Source of variation

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>740.03</td>
<td>1</td>
<td>0.64</td>
<td>3.84</td>
<td>0.423</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>177.63</td>
<td>1</td>
<td>0.15</td>
<td>3.84</td>
<td>0.694</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>433.20</td>
<td>1</td>
<td>0.38</td>
<td>3.84</td>
<td>0.540</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>135648.13</td>
<td>116</td>
<td>1151.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>119</td>
<td>1151.25</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Source of variation

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>869.41</td>
<td>1</td>
<td>0.75</td>
<td>3.84</td>
<td>0.386</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>57.41</td>
<td>1</td>
<td>0.05</td>
<td>3.84</td>
<td>0.824</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>120.00</td>
<td>1</td>
<td>0.10</td>
<td>3.84</td>
<td>0.748</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>136758.18</td>
<td>116</td>
<td>1158.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>119</td>
<td>1158.03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TPB Perceived behavioural control

**Source of variation**

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>858.67</td>
<td>1</td>
<td>0.73</td>
<td>3.84</td>
<td>0.392</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>14.70</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>0.911</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>165.67</td>
<td>1</td>
<td>0.14</td>
<td>3.84</td>
<td>0.707</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>138581.45</td>
<td>116</td>
<td>1173.28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>119</td>
<td>1173.28</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Source of variation**

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>4.80</td>
<td>1</td>
<td>0.00</td>
<td>3.84</td>
<td>0.949</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>1032.53</td>
<td>1</td>
<td>0.88</td>
<td>3.84</td>
<td>0.348</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>252.30</td>
<td>1</td>
<td>0.21</td>
<td>3.84</td>
<td>0.643</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>138453.87</td>
<td>116</td>
<td>1174.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>119</td>
<td>1174.32</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Source of variation**

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>1147.01</td>
<td>1</td>
<td>0.98</td>
<td>3.84</td>
<td>0.323</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>28.03</td>
<td>1</td>
<td>0.02</td>
<td>3.84</td>
<td>0.877</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>85.01</td>
<td>1</td>
<td>0.07</td>
<td>3.84</td>
<td>0.788</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>138250.45</td>
<td>116</td>
<td>1172.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>119</td>
<td>1172.36</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Lung cancer risk perception

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>49.41</td>
<td>1</td>
<td>0.04</td>
<td>3.84</td>
<td>0.837</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>1598.70</td>
<td>1</td>
<td>1.37</td>
<td>3.84</td>
<td>0.241</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>1222.41</td>
<td>1</td>
<td>1.05</td>
<td>3.84</td>
<td>0.305</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>135528.48</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>135535.73</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>2288.13</td>
<td>1</td>
<td>1.97</td>
<td>3.84</td>
<td>0.160</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>7.50</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>0.936</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>537.63</td>
<td>1</td>
<td>0.46</td>
<td>3.84</td>
<td>0.496</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>135353.73</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>135355.08</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>403.33</td>
<td>1</td>
<td>0.35</td>
<td>3.84</td>
<td>0.555</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>837.41</td>
<td>1</td>
<td>0.72</td>
<td>3.84</td>
<td>0.395</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>21.68</td>
<td>1</td>
<td>0.02</td>
<td>3.84</td>
<td>0.891</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>136415.08</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>136415.08</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>s2</th>
<th>Hcalc</th>
<th>Hcrit</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>2288.13</td>
<td>1</td>
<td>1.97</td>
<td>3.84</td>
<td>0.160</td>
<td></td>
</tr>
<tr>
<td>Rows</td>
<td>7.50</td>
<td>1</td>
<td>0.01</td>
<td>3.84</td>
<td>0.936</td>
<td></td>
</tr>
<tr>
<td>Interaction</td>
<td>537.63</td>
<td>1</td>
<td>0.46</td>
<td>3.84</td>
<td>0.496</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>135353.73</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>135355.08</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 25: Self-reported likelihood of seeing a doctor about symptoms (Trial 2)

Primary outcome intention measure (Trial 2)

How likely do you think it is that you will see a doctor about your symptoms in the next 3 weeks?

[Scale with options: 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%]
Appendix 26: Initial survey (Trial 2)

Initial survey (Trial 2)

Do you live in the UK?

☐ Yes
☐ No

What is your age?

☐ Under 40
☐ 40 and over

What is your gender?

☐ Male
☐ Female

Do you smoke?

☐ Yes, I currently smoke
☐ No, but I used to smoke
☐ No, I have never smoked

Do you have any of the following conditions? (select all that apply)

☐ asthma
☐ chronic obstructive pulmonary disease (COPD)
☐ cystic fibrosis
☐ pulmonary fibrosis
☐ a respiratory infection
☐ sleep apnoea
☐ tuberculosis (TB)
☐ other respiratory/lung conditions

☐ none of the above

Which of the following symptoms do you have?

(Tick all that apply)

☐ A cough

☐ A change in a long-standing cough

☐ Coughing up phlegm with blood in it, or spitting blood

☐ Feeling breathless

☐ Discomfort in the chest, shoulders or back

☐ A change in the voice (such as hoarseness, sounding deeper, weaker, or quivering)

☐ Noises when breathing (such as wheezing or harsh breath sounds)

☐ Unintentional weight loss

☐ Feeling more tired than usual

☐ Any swellings or lumps around the face or neck area

☐ Recurring or persistent chest infections

☐ Changes in the appearance of fingers or fingernails

Have you already seen a doctor about your symptoms?

☐ Yes

☐ No

Are you researching your own symptoms, or on behalf of someone else?

☐ Own symptoms

☐ Someone else
Appendix 27: Next steps measure (Trial 2)

We can help by pointing you to some other useful websites. What would you like to do now?

- Make an appointment with my GP
- Register with a GP
- Find an NHS walk-in centre
- None of these
I intend to see a doctor about these symptoms within the next 3 weeks.

For me to see a doctor about these symptoms within the next 3 weeks would be

Seeing a doctor about these symptoms within the next 3 weeks would waste the doctor’s time.

Wasting my doctor’s time is

If I see a doctor about these symptoms within the next 3 weeks, the doctor may find a serious cause.

Finding a serious cause is

Seeing a doctor about these symptoms within the next 3 weeks, even if they are unlikely to be caused by lung cancer, can help to put your mind at rest.
Putting your mind at rest is

Seeing a doctor about these symptoms within the next 3 weeks could help to detect lung cancer early and treat it more effectively.

Detecting lung cancer early and treating it more effectively is

Most people who are important to me would want me to see a doctor about these symptoms within the next 3 weeks.

My family and friends would want me to see a doctor about these symptoms within the next 3 weeks.

My doctors would want me to see a doctor about these symptoms within the next 3 weeks.

When it comes to matters of health, I want to do what my doctors think I should do.
When it comes to matters of health, I want to do what my family and friends think I should do.

I expect that my doctor's practice will not have many appointments free in the next 3 weeks.

My doctor's practice not having many appointments available will prevent me from seeing a doctor about these symptoms within the next 3 weeks.

For me to make an appointment to see a doctor about these symptoms would be
Appendix 29: Follow-up questionnaire (Trial 2)

The follow-up questionnaire asked users different questions depending on their responses. The different options are shown below.

1) A user who reports not having sought help and who reports still experiencing symptoms is asked whether they intend to seek help in future.

![Follow-up questionnaire](image)
2) A user who reports not having sought help and who reports no longer experiencing symptoms is not asked whether they intend to seek help in future.
3) A user who reports having sought help is asked whether they think that the information received through the study encouraged them to discuss symptoms with a doctor / make an appointment.
4) A user who reports discussing symptoms with a doctor is asked what the outcome of the consultation was.

Follow-up questions
Tell us how you have been doing since you took part in our study.

Please enter your email address:

Since you took part in our “When2Go” study about three weeks ago, have you contacted your doctor to make an appointment to have your symptoms checked?
- Yes
- No
- I have contacted my doctor, but for another reason

Have you discussed your symptoms with a doctor?
- Yes
- No

Are you still experiencing the symptoms?
- Yes
- No

What was the outcome of your consultation with your doctor?
- Medication was prescribed
- Referral for a chest x-ray
- No treatment or referral
- Other

Would you say that the information you received through our study encouraged you to discuss your symptoms with a doctor / make an appointment?
- Yes
- No

If you have any further comments, please tell us in the box below.
Participant Information Sheet (PIS) (Trial 2)

EVALUATION OF A TOOL TO HELP PEOPLE APPRAISE SYMPTOMS AND DECIDE WHEN TO SEEK MEDICAL HELP

Julia Mueller, PhD Candidate at the University of Manchester

You are being invited to take part in a research study as part of a PhD degree at the University of Manchester. Please take time to read the following information carefully and take time to decide whether you wish to take part.

What is the purpose of the study?

For certain symptoms that may be associated with more serious illnesses, such as lung cancer, it is especially important to provide clear and helpful information to help people decide when they need to see a doctor about their symptoms. We have therefore developed a tool to help people appraise their symptoms and to encourage people to see their GP if necessary. The aim of this study is to find out how people use this tool and whether it is effective. This tool is not designed to help you diagnose yourself. It will help you decide whether you need to see a GP about your symptoms. If this tool does suggest that you should see a GP, this does not necessarily mean that you have lung cancer. It just means it would be better to get your symptoms checked out.

Why have I been invited to take part in this study?

We are inviting anyone to participate who is above the age of 18, lives in the UK and is experiencing any of the symptoms listed below.

- A cough (whether dry and tickly or productive with phlegm)
- A long-standing cough that changes or gets worse
- Feeling out of breath
- Discomfort in the chest, shoulder or back
- Coughing up phlegm with a blood in it (even if just a few specks) / spitting blood
- Changes in your voice
- Unexplained weight loss or unexplained loss of appetite
- Swelling of face and neck
- Persistent / recurring chest infection
- Tiredness or lack of energy
- Any changes in the appearance of the fingers or fingernails (such as a softening of the nailbed, stronger than normal curving of the fingernails, or thickening of the fingertips so that the shape looks like an upside down spoon)

**Do I have to take part?**

No, you do not have to take part in this study if you do not want to. It is up to you to decide whether or not to take part. If you decide to take part you are still free to withdraw at any time without giving a reason.

**What will taking part involve?**

The study will take place online. If you decide to take part, you will first be asked a few basic questions about yourself, or about the person who has been experiencing the symptoms. You will then be guided through the tool, which will provide information on symptoms and risk factors. In the end, you will receive advice on whether you should make an appointment to see your GP. Finally, you will be asked to fill in a questionnaire. Taking part in this study will take approximately 15-20 minutes. You will not be able to save the questionnaire and return to it at a later time so you will need to complete it in one go. If you are willing to be contacted for some follow-up questions about three weeks after taking part in the study, we will send you an email with a link for a brief questionnaire (this will take 2-3 minutes to complete). If we do not hear back from you, we will send you one reminder. After that we will not contact you again. Participating in the follow-up is optional; you can still do the main study even if you do not wish to be contacted afterwards.

**What are the possible disadvantages and risks of taking part?**

It is not our aim to ask any questions or provide any information that might be upsetting or stressful to you. If there are any questions you do not wish to answer or you feel you want to withdraw from the study at any point, that is absolutely fine. If you feel distressed or worried after taking part in this study, it will most likely be best if you contact your GP and speak to your GP about your concerns. If you would like further support or advice, you can also call MacMillan Cancer support for free on 0808 808 00 00.

**What are the possible benefits of taking part?**

We cannot promise the research will help you but you may find the information provided here useful in deciding whether you need to see a GP about your symptoms, or whether you should encourage someone else to see a GP.

**How is confidentiality maintained?**
All data will be stored securely at the University of Manchester. Any information collected about you during this study will be kept strictly confidential. Data from the questionnaires will be made anonymous, so your name will not be included. If you provide your email address for the follow-up questionnaire, this will be stored in a separate database from the remaining data, on a secure server. All data will be stored in an encrypted format so they can only be accessed with the correct password, to ensure only the research team can access the data. After the study is completed, the dataset will be fully anonymised and any information that could make participants identifiable will be removed. The anonymised dataset will then be stored in a publically accessible repository, according to data management policy of the Medical Research Council. Data from the study will be kept for a minimum of 10 years after the date of any publication which is based upon it, to follow recommended good practice guidelines for research.

Study data and material may be looked at by individuals from the University of Manchester or from regulatory authorities, for monitoring and auditing purposes, and this may well include access to personal information.

**What will happen if I do not want to carry on with the study?**

You can withdraw from the study at any time without giving a reason. From the moment you withdraw, we will not collect any further data from you.

**What if I want to complain?**

If there is a problem, you may contact me (Julia) in the first instance or you can contact my supervisor Professor Chris Todd.

If there are any issues regarding this research that you would prefer not to discuss with me or my supervisor, please contact the Research Practice and Governance Co-ordinator by either writing to 'The Research Practice and Governance Co-ordinator, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester M13 9PL', by emailing: research.complaints@manchester.ac.uk, or by telephoning 0161 275 7583 or 275 8093.

**What will happen to the results of the study?**

The data obtained will be analysed and written up as part of a PhD thesis. The data will also be summarised in the form of research articles which will be published in a scientific journal.

**Who is funding the research?**

This research is funded by the Medical Research Council.
Who has reviewed the study?

Ethical approval for this study was granted by The University of Manchester Research Ethics committee (Ref #15353).

Whom can I contact for further information?

If you have any questions or require any additional information about this research project please do not hesitate to contact me or my supervisor at:

**Julia Mueller (chief investigator)**

*School of Computer Science,*  
*LF1 Kilburn Building,*  
*Oxford Road, MANCHESTER,*  
*M13 9PL, UK*

[julia.mueller@manchester.ac.uk](mailto:julia.mueller@manchester.ac.uk)  
Tel.: 0161 275 78 21

**Professor Chris Todd (supervisor)**  
**Professor of Primary Care and Community Health**

The School of Health Sciences  
University of Manchester  
Room 6.314b, Jean McFarlane Building  
Oxford Road  
Manchester M13 9PL  
[chris.todd@manchester.ac.uk](mailto:chris.todd@manchester.ac.uk)  
Tel.: 0161 306 7865

If you would like some additional information on lung cancer you can contact:

**Macmillan Cancer support**  
[www.macmillan.org.uk](http://www.macmillan.org.uk)  
Macmillan Cancer Support  
89 Albert Embankment, London, SE1 7UQ  
Phone: 0808 808 00 00

**The Roy Castle Lung Cancer Foundation**  
[www.roycastle.org](http://www.roycastle.org)  
The Roy Castle Centre (Head Office)  
Enterprise Way, Wavertree Technology Park  
Liverpool, Merseyside, L13 1FB  
Phone: 0333 323 7200  
Email: foundation@roycastle.org