Measuring the Impact of Using Health-Related Websites: The eHealth Impact Questionnaire

Laura Kelly

Nuffield Department of Population Health

Harris Manchester College, University of Oxford

A thesis submitted in Trinity Term 2014 to the Medical Sciences Division,
University of Oxford, for the degree of Doctor of Philosophy

Contents

Appendices

CHAPTER 9 APPENDICES

Tables

Figures

List of Abbreviations

Abbreviation	Definition
AIC	Anti-image correlation
ANOVA	Analysis of Variance
BME	Black and Minority Ethnic
BOS	Bristol Online Survey
CFA	Confirmatory Factor Analysis
CI	Confidence Interval
CJ	Professor Crispin Jenkinson
CSS	Cascading Style Sheets
CTT	Classical Test Theory
df	Degrees of freedom
DIF	Differential Item Functioning
DIPEx	Database of Individual Patient Experiences
EFA	Exploratory Factor Analysis
eHIQ	eHealth Impact Questionnaire
GP	General Practitioner
HERG	Health Experiences Research Group
HIV	Human Immunodeficiency Virus
HSRU	Health Services Research Unit
HTML	HyperText Markup Language
HTO	Health Talk Online
ICC (1)	Intra Class Coefficient
ICC (2)	Item Characteristic Curve
IP	Internet Protocol address
KMO	Kaiser-Meyer-Olkin statistics
LK	Laura Kelly
LTC	Long Term Conditions
MND	Motor Neurone Disease
MS	Multiple Sclerosis
MSA	Measure of sampling adequacy
NHS	National Health System
OII	Oxford Internet Institute
OXIS	Oxford Internet Survey
PCA	Principal Components Analysis
PCT	Primary Care Trust
PEx	Patient Experiences
PSI	Person Separation Index
RCT	Randomised Controlled Trial
RMT	Rasch Measurement Theory
SD	Standard Deviation
SF eHIQ	Short Form eHealth Questionnaire
SZ	Professor Sue Ziebland
UK	United Kingdom
US	United States
WHO	World Health Organisation

Acknowledgements

I would like to express my deepest appreciation to all those who made it possible for me to complete this thesis. First and foremost, I would like to thank my supervisors, Professor Crispin Jenkinson and Professor Sue Ziebland, for their unending support and advice throughout this doctoral research. I would particularly like to thank Crispin for being a constant source of support, reassurance and encouragement. I am grateful for his knack of keeping a sense of perspective and the faith he seemed to place in my abilities! Without Sue, the opportunity to complete this research would not have arisen. During the course of this research she has been a continuous source of enthusiasm and has really encouraged me to think carefully about the questions which needed to be addressed. I feel very fortunate to have had both supervisors' guidance and really appreciate the time they have spent in our many supervision sessions throughout the past few years. Thank you for keeping me focused on the task at hand!

I would like thank my funders, the National Institute for Health Research (NIHR), who made it possible for me to complete this research through their support of the iPEx Programme under its Programme Grants for Applied Research funding scheme (RP-PG-0608-10147).

This research would not have been possible without the help and input of the numerous stakeholders and research participants who advised and took part in this research. To them all, I am hugely grateful for their time and effort.

There are many others who helped me throughout this work. My thanks goes to everyone at the Health Services Research Unit and Health Experiences Research Group who not only made me feel very welcome when I arrived in Oxford, but also encouraged me to keep going when the going got tough! A huge thank you goes to my 'office-mates', Sarah Dummett and Paul Hewitson who provided much support and laughs each day. Special thanks also go to members of the iPEx

Programme for their interest in my research and feedback on methodology. My particular thanks go to Angela Martin for her administrative support, Fadhila Haeri Mazanderani for her advice and encouragement, the five iPEx members and Professor Vikki Entwistle who participated in my item pool expert panel, the Northumbrian iPEx colleagues for access to transcripts arising from focus groups which they had previously carried out and, finally, the iPEx trial group and Julie Evans (HERG) at Oxford who have been an enormous help through incorporating this work into their research.

I would also like to take this opportunity to thank my early mentors, Dr Pauline Conroy, Dr Sally Spencer and Dr Geraldine Barrett for their initial encouragement to apply for this doctoral position and their continued support.

Massive thanks go to my husband to be, Asher Anderson, who has been a constant source of support to me throughout the ups and downs of this thesis (there have been many and surprisingly he still wants to sign the contract!). I’d also like to thank his X-Box for keeping him company in my absence- the services of which are (hopefully) no longer required! Finally I would like to thank my parents, family and friends who have also been enormous supports to me long before I even stepped foot in Oxford!

Funding

This research was funded by a wider NIHR funded programme entitled: *Examining the role of information on patients' experiences as a resource for choice and decision-making in health care*. This Oxford-led programme is being carried out in collaboration with research groups in Warwick, Glasgow, Northumbria and Sheffield.

The iPEx programme presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research funding scheme (RP-PG-0608-10147). The views expressed in this programme are those of the authors, representing iPEx, and not necessarily those of the NHS, the NIHR or the Department of Health.

The iPEx study group is composed of:

- University of Oxford (Sue Ziebland, Louise Locock, John Powell, Andrew Farmer, Crispin Jenkinson, Rafael Perera, Ruth Sanders, Angela Martin, Laura Griffith, Susan Kirkpatrick, Nicolas Hughes, Laura Kelly, Braden O'Neill, Nikki Newhouse and Ally Naughten)
- University of Northumbria (Pamela Briggs, Elizabeth Sillence, Claire Hardy)
- University of Sussex (Peter Harris)
- University of Glasgow (Sally Wyke)
- University of Warwick (Fadhila Mazanderani)
- Department of Health (Robert Gann)
- Oxfordshire Primary Care Trust (Sula Wiltshire)
- User advisor (Margaret Booth)

Publications and presentations

The work presented in this thesis has been published and presented previously.

Paper:

Kelly, L., Jenkinson, C. and Ziebland, S. 2013. Measuring the effects of online health information for patients: Item generation for an e-health impact questionnaire. *Patient Education and Counseling*, 93, 433-438.

Oral presentations:

Kelly, L., Ziebland, S. and Jenkinson, C. 2013 Developing a tool to measure the effects of using health-related websites (the eHealth Impact Questionnaire). *20th Annual Conference of the International Society of Quality of Life Research,* Quality of Life Research, 22, 1-137. Presented at *ISOQOL, Miami, Florida. October 9-12,2013*

Kelly, L. 2012. Measuring the effects of exposure to health websites. Presented in symposium: Hardy, C., Hinton, L., Mazanderani, F. and Kelly, L. 2012. Communicating with other patients online: Exploring the effects of sharing health experiences on the internet. *EACH, International Conference for Communication in Healthcare, St Andrews, Scotland. 4-7 September 2012*

Kelly, L., Jenkinson, C. and Ziebland, S. 2012. Measuring the effects of online patient experience. Presented at *Work-in-Progress seminar, Department of Public Health, University of Oxford. 31.05.12*

Poster abstracts:

Kelly, L., Ziebland, S. and Jenkinson, C. 2013. The E-Health Impact Questionnaire: Developing a Tool to Measure the Effects of Using Health-Related Websites. *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 16, A605. Presented at *ISPOR 16th Annual European Congress, Dublin, 2-6 November 2013*

Kelly, L. 2013. Measuring the effects of using health-related websites: Translatability Assessment of questionnaire items in the eHealth Impact Questionnaire (eHIQ). Presented at *Medicine 2.0, 6th World Congress on Social Media, Mobile Apps, Internet/Web 2.0. September 23-24, 2013*

Kelly L., Simpson H., Verjee-Lorenz A., Clayson D., Churchman D., Jenkinson C. 2013. Translatability assessment of the e-health impact questionnaire (E-HIQ). *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research*, 16, A45. Presented at *ISPOR 18th Annual International Meeting, New Orleans, May 19-22, 2013*

Kelly, L., Ziebland, S. and Jenkinson, C. 2012. Developing and pre-testing an item pool relating to the effects of exposure to health websites. *19th Annual Conference of the International Society of Quality of Life Research.* Quality of Life Research, 21, 1-132. Presented at *ISOQOL, Budapest, October 24-27, 2012*

Measuring the Impact of Using Health-Related Websites: The eHealth Impact Questionnaire

Laura Kelly, Nuffield Department of Population Health and Harris Manchester College, University of Oxford

A thesis submitted in Trinity Term 2014 to the Medical Sciences Division, University of Oxford, for the degree of Doctor of Philosophy

Abstract

Introduction

Health-related websites have developed to be much more than information sites: they are used to exchange experiences and find support as well as information and advice. It is important that health professionals and website developers understand how content may impact users. This thesis documents the development and application of a tool to measure the impact of using health-related websites which contain experiential and/or factual information.

Methods

A multi-method study with five stages. Stage 1: Questionnaire items based upon themes relating to the impact of using health-related websites were constructed following qualitative secondary analysis of 93 interviews relating to patient and carer experiences of health and a recent literature review. Items were assessed by an expert panel. Stage 2: Cognitive interviews were carried out to confirm acceptability of items. Stage 3: Item reduction steps were used to reduce the number of items. Stage 4: The validity and reliability of the remaining items were tested using traditional and modern psychometric methods. Stage 5: The new questionnaire was piloted in a randomised controlled trial.

Results

Eighty-two items were constructed according to the key themes identified in Stage 1. Following expert and patient refinement, two independent item pools entered psychometric testing. The first item pool related to general views of using the internet in relation to health and the second item pool related to the impact of using a specific health-related website. Sub-scales and summary scores were found to have high construct validity, internal consistency and test-retest reliability. The questionnaire showed high completion rates and low counts of missing data in a trial setting.

Conclusion

Analysis confirmed good psychometric properties in the eHIQ-Part 1 (11 items) and the eHIQ-Part 2 (26 items). Preliminary findings of trial data demonstrate the acceptability and feasibility of including the eHIQ in randomised controlled trials. This tool will enable the measurement of the impact of health-related websites containing various styles of information and support across a range conditions and facilitate their accurate evaluation in clinical trials.

CHAPTER 1: Overview of the research

1.1 Introduction

The internet is a resource which has huge potential for promoting health and preventing disease (Reis et al., 2013). One important way in which the internet has been used in connection with health is through the distribution of information throughout both industrialised and developing nations (Siegel et al., 2009, Higgins et al., 2011). To date, however, there has been little research into the effects of using the web for health information. This thesis explores what is already known about the impact of using online health information and considers the difficulties of accurately measuring the effects of using various health-related websites. The development of a suitable tool with which to measure the impact of using health-related websites is subsequently documented and preliminary findings of comparisons between groups using prototype websites are outlined.

General use of the internet during the years 2000 to 2009 increased globally by almost 400% and became an integral part of everyday life for a substantial number of people (Higgins et al., 2011). With this rise, the internet has become integrated into everyday healthcare by millions of people. In the UK, for example, 73% of adults accessed the internet daily with people across all stages of life finding information online to improve health (students (32%), employed (45%) and retired (39%)) (ONS, 2013). This increase is, in part, due to increased ease of high speed internet access and the availability of a wider range of internet connection devices such as smart phones and other hand held tablet devices. Fifty-five percent of people aged 16 or over own a smart phone in the USA, while 51% in the same age group own a smart phone in the UK (Rainie and Smith, 2013, Ofcom, 2013). Similarly tablet use has increased with 35% ownership in the USA and 24% ownership in the UK (Ofcom, 2013, Rainie and Smith, 2013). The portability and ease of use associated with handheld devices have made instant access much easier when compared to operating a traditional desktop computer and can encourage those with low levels of computer skills to get online (Prainsack, 2013). In addition to improved internet access, the speed of connection is also becoming faster with 86% of households in the UK now having access to a broadband connection (Eurostat, 31.07.2013).

With 86% (43.8 million) of people aged 16 or above in the UK having used the internet, the web is undoubtedly a valuable resource with which both health professionals and ’health consumers’ can disseminate information and connect with each other (ONS, 2013 Table 1A, ONS, 2013, Table 1B). Figures indicate that the digital divide (i.e. the gap between those who have the ability and skills to access to the internet and those who do not) appears to be narrowing and this is somewhat indicated by the increased use of the internet amongst Europeans across all age groups (Kummervold et al., 2008). Unsurprisingly, younger people in the UK are still more likely to have used the internet compared to older people; however, approximately two thirds of people aged 65-75 have now reported using the internet (ONS, 2013, Table 1B). Just over two thirds of people with a disability[1] have used the internet and 95.2% of people with low earnings (less than £200 per week[2]) have also reported using the internet (ONS, 2013, Table 1B, ONS, 2013, Table 5B). Figures indicating a rise in internet use among those with a disability have been

---

[1] Self-assessed in line with the Disability Discrimination Act (DDA)

[2] £200 amounts to 31.7 hours per week working at minimum wage (£6.31)

further supported by Oxford Internet Survey (OxIS) data showing an 11% increase in use from 2011 to 2013 (40% to 51%) (Dutton et al., 2013). Overall, these figures indicate that access to the internet is penetrating key socio-demographic groups and that there is an increasing potential for information to reach all sections of society.

While it is relevant to acknowledge the rising number of people getting online, it is particularly important to this thesis to understand the extent to which health-related websites are accessed and the demographic make-up of their audience. The *Pew Internet and American Life Project* found that the third most popular reported activity on the internet was to look for health information (Fox, 2011a). Various estimates have been made regarding the number of people who use the internet to obtain health information. In 2013, it was estimated that 45% of UK individuals aged 16-74 had sought health-related information, slightly above the EU average of 44% (Eurostat, 23.01.2014). Oxford Internet Institute (OII) research however found that a considerably higher figure (71%) have reported sourcing health information online (Dutton and Blank, 2011). In addition to those accessing online information directly, many people have accessed the internet through a family member and/or friend looking for information on their behalf (Ayantunde et al., 2007) and approximately one-quarter of people in the UK who use the internet report accessing and/or posting their own health information online (O'Neill et al., 2014).

Information can be sought for a range of health issues. Research has shown one in five internet users went online to find people who have similar health concerns to them, with the number rising for those with a chronic condition (Fox, 2011b). Peer-to-peer contact has also been sought by caregivers, those experiencing a medical crisis and health behaviour groups (for example, those hoping to make changes to their weight or stop smoking) (Fox, 2011b). Estimating an accurate demographic profile of those using the internet in relation to health is difficult. Typically it is thought the people most commonly seeking health information are female, young, white, well-educated and high earners (Fox and Duggan, 2013, Ybarra and Suman, 2006, Atkinson et al., 2009). These characteristics do however appear to fluctuate. For example, no difference was found between internet use and educational qualifications among patients attending Obstetrics and Gynaecology clinics (Sethuram and Weerakkody, 2010). Previous research also indicated older people are just as likely to access a computer-based health information database containing detailed information on topics such as wellness, prevention, symptom identification, self-care and advice on when to seek help from healthcare professionals as compared to younger people (Wagner and Wagner, 2003). The latter may be interpreted somewhat cautiously as an online database may be considered to have a higher degree of regulation compared to freely searching online. Nonetheless, it does indicate that older people are willing to use the internet in relation to health.

1.2 Research context

It is clear that a substantial number of people, spanning a range of socio-demographic characteristics, are frequently using the internet and, of relevance to this thesis, are using it as a resource for accessing health information. Given their global and local reach, health-related websites may therefore be a hugely valuable resource in public health. The impact of online information and support however is largely unknown. A review has shown 70% of studies assessing online information concluded that information quality (in relation to accuracy, completeness, technical criteria, readability and design) was problematic (Eysenbach et al., 2002). The manner in which people use information gained from the web and the outcome of this process is therefore critical when forming future health information strategies.

Research has previously focused on comparing online information with other forms of information delivery (for example, leaflets or pamphlets) (Wantland et al., 2004). Using the web to source information compared to using printed materials however is a very different user experience and finding information which is comparable to an individual's cultural views or demographic profile is arguably more likely to occur online (NCB, 2010). The mass of websites available, their ever changing dynamic and their popularity suggests online health information should now be placed and assessed in its own genre. To do this, websites need to be compared and assessed using appropriate methods to fully understand how their content may have a positive or negative impact on users.

There is currently no suitable instrument with which to assess the impact of specific health-related websites. This thesis therefore aims to advance this area of research by examining the potential range of effects of using health-related websites and develop an instrument with which to capture these effects. Such a tool needs to enable websites to be compared in a standardised manner and have good psychometric properties. Developing a valid and reliable tool would advance the assessment and comparisons of health-related websites and help to inform health professionals, researchers and web developers about the value of including particular styles of information (for example, the value of including experiential information) on their websites.

1.3 Research aims

This thesis documents research which aimed to develop an instrument (The eHealth Impact Questionnaire) to capture the impact of using health-related websites. The term 'impact' refers to a number of influences a website might have upon those viewing it. Previous qualitative and intervention studies support the suggestion that there are many potential positive and negative effects of using online health information, particularly in relation to websites containing experiential information and support (Ziebland and Wyke, 2011). These effects can include gaining an understanding of health or becoming confused by misleading information, reducing or affirming feelings of support, changes in relationships with others through having an outlet to vent frustrations or over reliance on the 'virtual world' and changes in experiencing of health services. Further positive and negative consequences included people learning to tell their own story about health or being more adept at manufacturing accounts of health, becoming motivated by others to engage in healthy or unhealthy behaviour and the use of images to deliver important messages, but which may also cause a person to become anxious about their future health. These themes are explored more closely in the following chapters, however, for the purposes of this research, 'impact' refers to the extent to which using a website may influence some or all of these areas. This instrument needed to be suitable for use across a range of health conditions and various types of health-related websites. The instrument was developed throughout a number of stages. ***Stage One*** aimed to construct a set of meaningful and psychologically relevant items which relate to the impact of using health-related websites. ***Stage Two*** aimed to design and pre-test the online pilot questionnaire. ***Stage Three*** aimed to administer the pilot online questionnaire across a range of health conditions with a view to reducing and refining items. ***Stage Four*** aimed to evaluate the validity, reliability and scoring of the questionnaire using a combination of traditional and modern psychometric techniques**.** ***Stage Five*** aimed to pilot test the new questionnaire in a setting for which it is designed and provide preliminary results on the impact of using a prototype health-related website.

1.4 Thesis structure

This thesis is made up of a further ten chapters. Chapter Two, ***Health-related websites: technology and content,*** discusses the evolving and interactive internet landscape for today's health consumer. The types of information currently available online are discussed and an overview is given of what is known about health consumer's views of quality, trust and design elements of various websites. The chapter also outlines what is currently known about the impact of online health information and support and introduces the difficulties for measuring such impacts.

Chapter Three, ***The measurement of the impact of health-related websites: a structured review***, investigates how previous trials have assessed the impact of health-related websites using standardised measures. This review was carried out to confirm the lack of suitable measures for use in randomised controlled trials and, where no gold standard measure existed, to identify any potential reference measures which may be suitable for use in the validation of a new measure.

Chapter Four, ***The development of the eHealth Impact Questionnaire item pool***, outlines the research design, methodology and analysis used to develop an item pool relevant to the impact of using health-related websites. The chapter describes how items were primarily informed through a relevant literature review and secondary qualitative analysis of 93 narrative interviews relating to patient and carer experiences of health. Statements relating to relevant themes found through the literature and qualitative analysis were taken from interview transcripts and re-cast as questionnaire items. These questionnaire items were subsequently shown to an expert panel for review.

Chapter Five, ***The pilot eHealth Impact Questionnaire: User testing and web survey design***, describes how the item pool was arranged into an online questionnaire format (the pilot eHIQ) using appropriate survey software. The chapter subsequently describes the methods used to pre-test the initial questionnaire among patients and carers to improve face and content validity. A detailed account of the modification process is outlined during three rounds of cognitive debrief interviewing.

Chapter Six, ***The pilot eHealth Impact Questionnaire: Item reduction and scale generation,*** gives an account of the research design and methods used to reduce and refine the questionnaire resulting from cognitive debrief interviews. The chapter gives a detailed description of the statistical procedures used, including a description of the a-priori item reduction decision steps adopted to reduce the number of items within the questionnaire.

Chapter Seven, ***The eHealth Impact Questionnaire: Construct validity test-retest reliability and translatability***, gives an account of the research design and methods used to assess the psychometric properties of the questionnaire. Reliability and validity of the new questionnaire are tested and the final sub-scales are identified. A scoring algorithm is presented for each sub-scale identified.

Chapter Eight, ***The eHealth Impact Questionnaire: Summary score and short form***, outlines the methodology used to carry out a higher order factor analysis and the development of a summary index score for the eHIQ-Part 1 and e-HIQ Part 2. A short form version of the questionnaire is also developed for use in studies which require fewer items.

Chapter Nine, ***Creating interval level scaling for the eHealth Impact Questionnaire***, outlines the methodology used to explore the feasibility of creating equal interval level measurement for each sub-scale using Rasch measurement theory. Rasch based person location scores are subsequently contrasted with statistical findings resulting from scores derived from the scoring algorithm based on Classical Test Theory.

Chapter Ten, ***Comparing prototype websites using the eHealth Impact Questionnaire***, uses data from a randomised controlled trial comparing a prototype website with a control website. The websites shared design and multimedia features, but differed in website content whereby the active website contained scientific (or ‘factual’) information plus experiential information and the comparator website contained scientific information only. Secondary data analysis is conducted to provide preliminary descriptive statistics of the study and to explore any differences between groups using both Classical Test Theory and Rasch based scoring.

Chapter Eleven, ***Discussion and Conclusions,*** discusses the overall performance of the measures’ psychometric properties. Recommendations are made on how to use the eHealth Impact Questionnaire in the appropriate research setting. Challenges associated with recruiting for online surveys are discussed based upon lessons learnt in this research. The final chapter also includes concluding remarks on the overall findings of this thesis.

CHAPTER 2: Health-related websites: technology and content

2.1 Health-related websites and advances in web technology

Health-related websites are developed by a wide spectrum of information providers ranging from funded organisations (commercial, public and volunteer sectors) to providers with no financial backing (for example, private individuals creating websites or writing blogs). Emerging web technologies have various implications for how information is used and the nature of information that is available on these websites. Websites have conventionally presented medical information in the form of scientific facts with little reference to the experience of living with a health condition. Personal experiences however can be useful in maintaining a web user's interest, can provide more in depth information and provide opportunities to compare and contrast experiences of health (Sillence et al., 2007b). Such experiential information is now routinely incorporated into mainstream websites and stories about other people's health can currently be accessed on the UK National Health Services biggest health website (NHS Choices), national and local charitable groups, and private company websites.

Advances in web development have inevitably influenced how websites are used in relation to health. At first, the internet had a top-down nature meaning web users could view information, send emails and post on relatively simple text based forums. Today this period of internet use is referred to as 'Web 1.0'. The present era, referred to as Web 2.0 facilitates sharing, collaboration and connecting people and has resulted in websites including (or consisting entirely of) user-generated content (NCB, 2010). Many of the features which have contributed to the changing Web 2.0 landscape in health are encompassed in Box 2.1.

Box 2.1: Five themes which are relevant to Web 2.0 and health

Eysenbach (2008) identifies five themes connected to health which are relevant to Web 2.0. These are:

1) Social networking: Connections through a complex network of relations facilitating collaboration and filtering,

2) Participation: For example, maintaining personal health records,

3) Apomediation: In contrast to relying on traditional experts and authorities (gatekeepers) for access to information, ‘apomediation’ refers to an information seeking strategy where a person is guided to high quality information using various ‘apomediatory’ resources, such as lay-people (for example, other internet users and friends) and tools (for example, additional credibility cues and consumer ratings). These resources are not a requirement to finding information (as would be the case if information was provided by a health professional), but act as a filtering process to navigate the vast quantities of information available online

4) Collaboration: Bringing together groups who have not previously interacted such as scientists or user groups, and

5) Openness: Technical openness (drive for accessing and control of own data, and societal openness) and ability to access data which was previously not available to general public, for example, peer review journals.

In addition to websites presenting relatively static pages of information, Web 2.0 technology has made it easier to interact with others through platforms such as social networking forums and to use alternative ways of learning about health (for example, through reflecting on personal blogs and using multimedia). In health, multimedia use can facilitate interaction through text, audio, images and videos which may impact the interpretation of a message. Presenting and sharing information in such a manner has huge implications for the evolving landscape in health. Understanding the impact of these emerging technologies is key to integrating them effectively into healthcare (Pulman, 2010).

The variability in the way in which information is accessed and the various styles of information available on websites makes it is difficult to establish the positive and negative effects they may have on the user. Web users are likely to explore multiple websites in addition to offline resources when sourcing information (for example, in preparation for medical consultations) making the isolation of website specific effects even more challenging (Hu et al., 2012). Patterns of information seeking can also fluctuate. Web users may be influenced immediately or influenced over a period of time with searching patterns reducing or ceasing completely once a health decision has been made (Sillence and Briggs, 2009). A US national survey has also reported that people find information from health professionals most helpful for technical issues, but information and support from non-health professionals (for example, other patients) can be most helpful for issues related to coping and quick relief (Fox, 2011b). This suggests individuals value both professional and peer-to-peer advice at various points in their health journey and the manner in which they use the internet can vary accordingly. This suggestion is supported through qualitative analysis which found participants liked to underpin decisions with factual information, but valued experiential information to support these decisions in a variety of ways (for example, to recognise decisions, identify options, appraise options and support coping) (Entwistle et al., 2011).

The importance of establishing the impact health-related websites may have on the web user is critical in an evidence-based healthcare environment. Little work however has been carried out to identify or assess health outcomes in this area (Powell et al., 2005). Consideration, for example, should be given to whether conventional static ‘facts and figures’ provide the most useful forms of online information or whether the inclusion of other people’s experiences of health provide a more, or equally, informative and supportive experience. Influences and preferences of different styles of information should also be further researched so that information can be better targeted (Wyke et al., 2011).

Web developers and health care providers need to understand the potential effects of the information provided on their websites, whether it is provided directly by them or posted by a lay person. Examining the potential effects must also include understanding the effects of internet discussion forums, personal blogs, feedback and multimedia platforms may have on users. Assessing the impact of current forms of information technology can be challenging due to the time lag between technology advancement and evaluation (Siegel et al., 2009). However, knowing the impact of various styles of information is particularly important for the National Health Service (NHS) and web developers as they invest public resources and time in online information platforms. Exploring such issues may enable current technology to become integrated into healthcare using a patient perspective (Pulman, 2010).

2.2 Evaluating health-related websites

Various topics have been at the forefront of online health information research to date and should be drawn upon to understand what is known about how the internet is used in relation to health. Ethical and regulatory debates have centred on the protection of patients from inaccurate information (Prainsack, 2013). Credibility and trust markers on health-related websites were therefore encouraged as they were thought to be useful to web users when searching for information (Courtenay-Quirk et al., 2010). There have been many attempts to find suitable quality signposts for users. For example, in 2009, NHS England commissioned *The Information Standard* designed to certify information had been put through a process of quality assessment (including checks for clarity, accuracy, balance, based on current evidence) (NHS England, 2013b).

Given the heavy emphasis on regulation, many tools with which to assess quality of information have been developed. These tools use various codes of conduct and can be placed under the following groups: 1) Self-regulation or self-governance based code of ethics, 2) Rating tools, 3) Third party reviewers, 4) Accreditation systems; and 5) Certification systems (Hanif et al., 2009). Despite the range available however, many quality rating scales are not considered user friendly (for instance, many act in order to evaluate websites for awards and kite marks and are not ordinarily used by the lay person) and none have been universally accepted for health information evaluation (Bernstam et al., 2005, Hanif et al., 2009).

While several evaluative criteria consistently feature in quality rating tools, they may be considered somewhat arbitrary given quality is often dependent on the individual's perspective which may not be based upon evidence (Wilson and Risk, 2002). Quality tools can therefore fail to take into account that the lay user may assess quality in a different manner than health experts. For example, a lay user may rate information as high quality if they find it useful and are reassured, but the same information may be considered inaccurate by a health professional (Bernstam et al., 2005, Ademiluyi et al., 2003). Furthermore, research has shown that high credibility sources (for example, those which have been written by an expert, contain current information and have no conflicts of interest) compared to low credibility sources can have no difference on quality perceptions of the consumer (Bates et al., 2006).

The perceived trustworthiness of information subsequently featured heavily in research relating to health-related websites and has been found to predict a person's perception of the usefulness of the internet as an information source (Rains, 2007). Despite being the most common first source of information, it is not considered the most trusted information source (Marrie et al., 2013). Literature suggests that information is processed rapidly and websites are rejected on subjective quality criteria. For example, women searching for information on Hormone Replacement Therapy rapidly dismissed websites on the basis of (distrusting) poor design, advertising and the general feel of the website (Sillence et al., 2007b). On engaging with a website, the women in the same study assessed credibility in more depth through judging expertise, bias and authorship (in some cases distrusting a high quality website from a pharmaceutical company) (Sillence et al., 2007b).

Trust associated with using the internet is not static. The more interaction a person has with the internet, the more likely they are to trust it as a resource and develop the ability to deal with any associated risks (i.e. trust increases with user experience) (Dutton and Shepherd, 2006). Recent data also suggests that negative experiences of using the internet are expected by today's web users and, through having these expectations, trust does therefore not diminish with negative encounters (Blank and Dutton, 2012). These observations reflect the increasingly refined internet user with the ability to identify or 'weed out' poor information. It may be noted however that this analysis conducted by the Oxford Internet Institute (OII) focused upon e-commerce which may or may not hold true for negative online experiences in relation to health. Nonetheless it does hold a promising insight into how internet users are becoming increasingly accustomed to navigating the online environment (Blank and Dutton, 2012).

Quality, credibility and trust are important elements to be considered when exploring the use of health-related websites. However, while these elements are important in informing us about the factors leading to engagement with information, they do not inform us on how website content impacts on the individual. This thesis does not seek to re-examine factors associated with quality of information, but seeks to explore the effects of using specific health-related websites in the Web 2.0 environment. Identifying and understanding these effects is increasingly important considering the vast changes in the delivery and style of information in recent years.

2.3 What is known about the impact of website content?

Many consequences of using the internet to source health information have been highlighted in relation to their impact on health services. Research has shown patients use the internet to prepare, complement, validate and challenge information in the medical consultation (Caiata-Zufferey et al., 2010). Information can empower the patient and alter the traditional balance of power in the doctor-patient relationship (Pulman, 2010). Empowerment however should not necessarily be viewed as adversarial towards the health professional. Poor relationships with healthcare professionals have been found to be an unlikely motivator for seeking online information before a consultation, however, information sourcing can affect the discourse of a consultation (for example, the patient may query information sourced online) (Hu et al., 2012). Patients can therefore use the internet before visiting a health professional to make the best use of their consultation, show they are interested in their own health and to be taken seriously by the General Practitioner (GP) (Bowes et al., 2012). Pregnant women have also reported using the internet to increase knowledge, add to information provided by health professionals, check symptoms and have more control over decisions (Lagan et al., 2010).

As discussed earlier, many forms of information are available online today. Both scientific facts and experiential information are valued when using health services (Wyke et al., 2011). Descriptive survey data have indicated positive effects of peer-to-peer interactions within online communities include becoming better prepared for medical appointments, having a better understanding of treatments, higher involvement in healthcare decisions, improved communication with healthcare professionals and help in making decisions to consult a healthcare professional (Wicks et al., 2010, Seçkin, 2011). Within Seckin's study (2011), overall information and decisional support was also demonstrated among people with cancer using the cancer specific *Cyber Info-Decisional Empowerment Scale*. Such evidence is not generalizable beyond the cancer setting, and the instrument not appropriate for other conditions. In the case of Wick's study (2010), while it provides preliminary evidence that online experiential information and support is beneficial, it is limited by the descriptive nature of the items. Both studies had relatively low response rates and a very specific population sample where the typical participant was a frequent online support group or electronic platform member. As members of their respective online communities, they may have given favourable results (i.e. they may have been invested in their online community and be highly involved in their care).

Finding out about other people's accounts of illness does not only benefit the web user in the practical sense of understanding and learning about their illness, it can also foster a sense of social support and shared identity (Sillence et al., 2007b). Learning how similar patients and carers cope and manage their condition can also create a sense of empowerment (van Uden-Kraan et al., 2008, Mo and Coulson, 2010). In-depth discussions around health can arise through finding out about other people’s experiences and may prompt people to seek advice, give advice and form relationships with people like them (Frost and Massagli, 2008).

Online information dissemination is not solely text based and including images on websites can aid information retention (Frisch, 2013). There are also numerous ways in which people can gain and share information through visual or multimedia clips. Youtube, for example, enables people to post before and after video clips regarding their experiences of treatment and can promote advocacy by putting pressure on policy makers to make treatments available (O'Neill et al., 2012). Media clips may also be used by health professionals to inform potential recipients of treatment procedures. Giving people access to online visual resources of medical procedures for example may affect health-related decisions, particularly in relation to sensitive topics.

While there are many indications that online health information promotes positive outcomes, there are some concerns. Accessing information can also have detrimental effects on an individual’s health and some have suggested that the internet may not be a suitable information medium for specific population groups, for example, those with an eating disorder (Rouleau and von Ranson, 2011). Distressing images of advanced disease (such as an incurable degenerative condition like motor neurone disease) can cause emotional distress for people with the same condition (Mazanderani et al., 2012). Online information can also encourage people to make decisions which are not in their best interest. Unbalanced (or unrepresentative) narrative accounts of adverse events connected to vaccinations may unnecessarily increase a person’s perception of risk (Betsch et al., 2011).

The mass of information available online can be challenging for the web user. This is particularly relevant in the era of Web 2.0 where content can have little regulation or monitoring, for example, in online support groups where a high volume of irrelevant information and off-topic communication may appear (Pulman, 2010). High preferences for social interaction in online support groups have also been associated with low satisfaction in offline support (such as family and friends) (Chung, 2013). Using online support groups may therefore have positive consequences, such as addressing a need for support, however, negative consequences my include an overreliance causing isolation from the offline world.

A further concern is that use of the internet may promote negative behaviour in some health groups. However, while some websites have been found to promote extreme negative consequences (for example, facilitating suicide pacts), mental health conditions can also be benefited through encouragement to seek support, treatment and by helping people to cope (Biddle et al., 2008).

Due to the quantity and variable styles of information available, it is important to understand their various impacts on the user. Isolating the effects (or added value) of including various styles of information on a website however is difficult when using existing research. Some preliminary research has been collated in a recent comprehensive conceptual literature review which identifies seven potential domains as relevant to the effects of exposure to online experiential information (Ziebland and Wyke, 2012). The review suggests the effects of using online experiential information centre around feeling supported, having an increased understanding or knowledge about health, changes to online and offline relationships, assistance in navigating health services, learning the art of conveying health stories, the effects of the visual aspects of illness and changes to health behaviour (Ziebland and Wyke, 2012). Potential negative effects of using online experiential information were also identified for each of the seven domains. Poor, unbalanced and unregulated information, for example, may lead to worse decisions. Instead of feeling supported, other people's accounts of illness may actually increase anxiety or lead to unrealistic hopes or expectation. When considering how relationships with others may be effected, 'offline' relationships may suffer if a person becomes isolated from their everyday face-to-face relationships, for example, if they think only people they interact with online with can understand them. While some may experience an improved experience of service use, others may become frustrated if their raised expectations are not met, some may overuse services and some may become weary of using services on reading about negative experiences. The review also highlighted how people may learn to tell their own story in relation to their health, yet some may also learn how to embellish or fabricate an account of health. In relation to health behaviour, some may become motivated by others to engage in healthy activities, but others may become encouraged to engage in unhealthy behaviour. Finally, the use of images may make messages more explicit or memorable, however, they also have the potential to raise anxiety about their future health. In summary, this paper goes some way to identifying the ‘added value’ experiential information may have when exploring the effects of online information, yet it is also mindful that there are potential negative effects of its use. However, without a tool with which to measure these effects in a controlled environment, (for example in a randomised controlled trial) conclusions may be limited in their application to the general population.

2.4 How can the impact of website content be measured?

A number of studies have shown promise in capturing improved health outcomes for web based interventions. For instance, a review comparing web-based interventions to non-web based interventions showed improved outcomes in behavioural change, knowledge, participation in healthcare, body shape perception, health decline and weight loss maintenance (Wantland et al., 2004). Interactive health communication applications (i.e. computer-based packages which combine the provision of health information with at least one other additional interactive service such as decision or peer support) have also been found to have positive effects on knowledge, social support, behavioural and clinical outcomes, however little evidence exists to support their effects on use of services (Murray et al., 2005). Behavioural change studies have proved popular in current research, for example, through web based interventions for reducing cold and flu transmission (Yardley et al., 2011) and through tailored interventions spanning nutrition and diet, physical activity, alcoholism, smoking cessation, encopresis, eating disorders and general risk behaviours (Lustria et al., 2009). A recent review on Web 2.0 and self-management in chronic diseases concluded that Web 2.0 use is likely to be associated with positive behavioural and clinical outcomes, however had negligible effect on adherence to medication, biological outcomes and use of health care services (Stellefson et al., 2013). The unique feature of web based information is that it can be used solely for information seeking, or people may choose to interact with others, for example, in an online discussion forum. Therefore it is uniquely placed for individuals to select the extent to which they engage with the information provided.

While there are indications of positive outcomes in connection to using the internet for health information, high quality research needs to be carried out to support early findings. The potential outcomes of using specific health-related websites are difficult to accurately measure, particularly due to the range of information available online. Many relevant outcomes which are expected to be affected through the use of a specific website may be measured through existing generic or condition specific patient reported outcome measures (PROMs) or other psychological assessments. For example, anxiety can be measured using the State-Trait Anxiety Inventory (Spielberger CD et al., 1989) or patient empowerment can be assessed using the Patient Activation Measure (Hibbard et al., 2004). Such scales however, were not designed for this research context and are unlikely to be sensitive in this setting. A 12 month study, for example, which investigated the effects of computer and internet services on the quality of life of older adults, used a range of physical, social, emotional wellbeing, activity and autonomy measures. The study found no clear change on any instrument and concluded that it is hard to measure subtle changes with current measures (Slegers et al., 2008).

2.5 Conclusion

There is currently a need for a suitable instrument with which to accurately assess the impact of the various forms of online information and support available today. While quality and trustworthiness of information are important in website development there is a need for a new instrument which encapsulates the range of experiences and effects of browsing or interacting with a website.

A new tool would provide the opportunity to compare the potential impact different health-related websites may have on the web user. To adequately measure the impact of website use today, the instrument should be suitable for use in the Web 2.0 environment (i.e. be suitable for use among websites with scientific ‘facts and figures’, websites which contain experiential information and forums). It is also desirable that such an instrument would be suitable for use across conditions (i.e. generic health and patients or carers). As such, comparisons of the impact of websites for different health conditions could be carried out and results may inform allocations of resources.

CHAPTER 3: Measuring the impact of health-related websites: a structured review

3.1 Introduction

The use of the internet for health information and support has risen rapidly, yet the impact on health of using online information remains unclear. Research is needed to document the impact of internet use on health and to try to determine ways of maximising positive effects whilst limiting potential negative consequences. To do this, a suitable tool with which to measure the outcomes of using health-related websites needs to be available. Such a tool is needed to go beyond technical and design issues and, instead, aim to assess the impact of health-related website content.

The research context in which online information should be evaluated has also changed. Comparisons between traditional methods of information dissemination (for example, paper based or through a health professional) and electronic modes are somewhat outdated as these approaches offer very different user experiences. A suitable tool should be sensitive to comparisons between websites which contain different forms of information. Such comparisons, for example, may take the form of websites which contain scientific information only versus websites which contain scientific information plus experiential information.

Taking account of the above considerations, a structured review was carried out to identify any generic instruments which may be used to measure the impact of health-related websites. Due to the breadth of papers available and the desire for an instrument to be suitable for use in the context of a randomised control trial, this review was confined to controlled comparison studies which assessed standalone websites (i.e.

website only interventions as opposed to website plus, for example, counsellor support or paper based information). This review served as a preliminary check for existing measures before embarking on the development of a new instrument.

3.2 Methodology

This structured review aimed to identify instruments used in controlled comparison studies of two or more health-related websites. A literature search and several stages of screening were carried out. Studies were selected for further reading according to specific inclusion and exclusion criteria outlined in section 3.2.2. In the context of this review 'health-related website' referred to websites which contain scientific (factual) information or experiential health information. In order to focus the review, purely interactive websites (such as websites consisting entirely of discussion boards or forums) were not included. 'Impact' referred to outcomes or experiences directly relating to the use of a health-related website. Although user feedback relating to interface design is important in web development, in the context of this review, studies which solely examined design were not included. 'Controlled comparison studies' referred to both randomised controlled trials (RCTs) and non-randomised trials which compared two or more health-related websites in order to assess the effects of the intervention.

3.2.1 Search strategy

A literature search was conducted using eight electronic databases (Medline, Embase, Cinahl, Electronic and communications, Abstracts, The Cochrane library, Psyc Info, CSI) to identify studies assessing the impact of using health-related websites. Studies were restricted to the English language (between January 1990 and November 2011). The search terms used in this review were: Health information OR health education OR advice OR self-help OR (support adj5 (internet or World Wide Web or ehealth or e-health or online or on-line or website or cyber or www)) AND World Wide Web OR internet OR online OR on-line OR ehealth OR e-health OR website OR cyber OR www AND Effect* OR evaluat* OR measure OR tool OR instrument. Reference lists of relevant papers were also hand searched and experts in the field were consulted. Experts comprised of colleagues in the National Institute for Health Research, Programme Grants for Applied Research, RP-PG-0608-10147. Entitled: Examining the role of patients' experiences as a resource for choice and decision-making in health care.

3.2.2 Selection of studies

Studies which compared standalone health-related websites (i.e. website only interventions) using standardised psychometric measures were selected.

3.2.2.1 Screening

Abstracts were uploaded into Endnote reference manager (Thomson_Reuters, 1998-2010). The number of papers for review was condensed through reading paper titles followed by a period of reviewing remaining abstracts. Remaining papers were reviewed in-depth to determine their suitability for inclusion.

Studies were selected for further reading using the following inclusion criteria: 1) Studies which included an instrument to assess the effects of using a website containing health information, 2) The instrument used was suitable for self-completion or completion by proxy, 3) The instrument was available for completion by either hard copy or electronic modes, and 4) Participants were eighteen years or older. Studies which assessed interventions with biomedical or clinical indicators (for example, blood pressure) only and studies assessing interventions which went beyond the scope of browsing a website (for example, online Cognitive Behavioural Therapy interventions, online health programmes or modules) were excluded. Computer training interventions and qualitative studies were also excluded for the purposes of this review.

3.2.2.2 Instrument assessment criteria

Key assessment criteria relating to the reliability and validity of the instruments were used in this review. Internal reliability (or internal consistency) indicates a scale is free from random error and that responses to similar items are highly correlated (Norman and Streiner, 2000). A reliable measurement tool should also demonstrate stability of scores over time (test-retest reliability) (Kline, 2000, Reeve et al., 2013, FDA, 2009). Test-retest reliability stipulates that, providing there has been no underlying change in the construct being measured, an instrument should produce the same (or very similar) scores when completed by a participant on two or more occasions (Jenkinson and McGee, 1997, Hays and Revicki, 2005).

Validity must also be demonstrated when assessing the quality of an instrument. Content and face validity examine whether items appear to measure the purported concept while construct validity assesses the evidence that test scores can be said to measure the intended construct (Messick, 1989). Evidence of construct validity can be shown through concurrent validity (moderate correlations with an identified measure which is hypothesised to be related to the construct being measured), predictive validity (correlation with future results in a similar construct) or criterion validity (correlations with a gold standard measure assessing similar domains) (Messick, 1989, Streiner and Norman, 2008). See Table 3-1 for apriori quality indicators used.

Table 3-1 Quality criteria for identified measures

Criteria	Indicator
Reliability	
Internal consistency	Cronbach's α >0.7 Pearson Separation Index > 0.7
Test- retest	Intra Class Correlation Coefficient >0.7
Validity	
Face and Content	All domains represented by items and tested with users
Criterion	Correlations ranging from 0.40-0.70 with 'gold standard'
Construct	Evidence of discriminant or convergent validity Known group differences

3.3 Results

Of the 11,238 studies identified in the initial word search, 42 were deemed suitable for in-depth screening (see Figure 3-10). Of the 42 papers identified, 14 papers were systematic review studies. These review papers were screened to ensure no relevant studies had been missed in the initial search. No further studies found in the review papers met the inclusion criteria for this review. Following in-depth screening of the remaining 28 papers, one study met the inclusion criteria. One further paper was included as of interest due to the methodological approach taken and instruments used (see Figure 3-1). Therefore, in total, two papers were selected for inclusion in this structured review. In each paper, two instruments were identified and assessed. This meant a total of four instruments were reviewed. Table 3-2 gives a synopsis of the relevant instruments included in the papers reviewed. Many of the excluded papers were not relevant to this review as they referred to the effects of using health-related websites, but did not use any standardised instruments designed for measuring the impact of a website intervention.

Figure 3-1 Flow diagram for selection of included studies

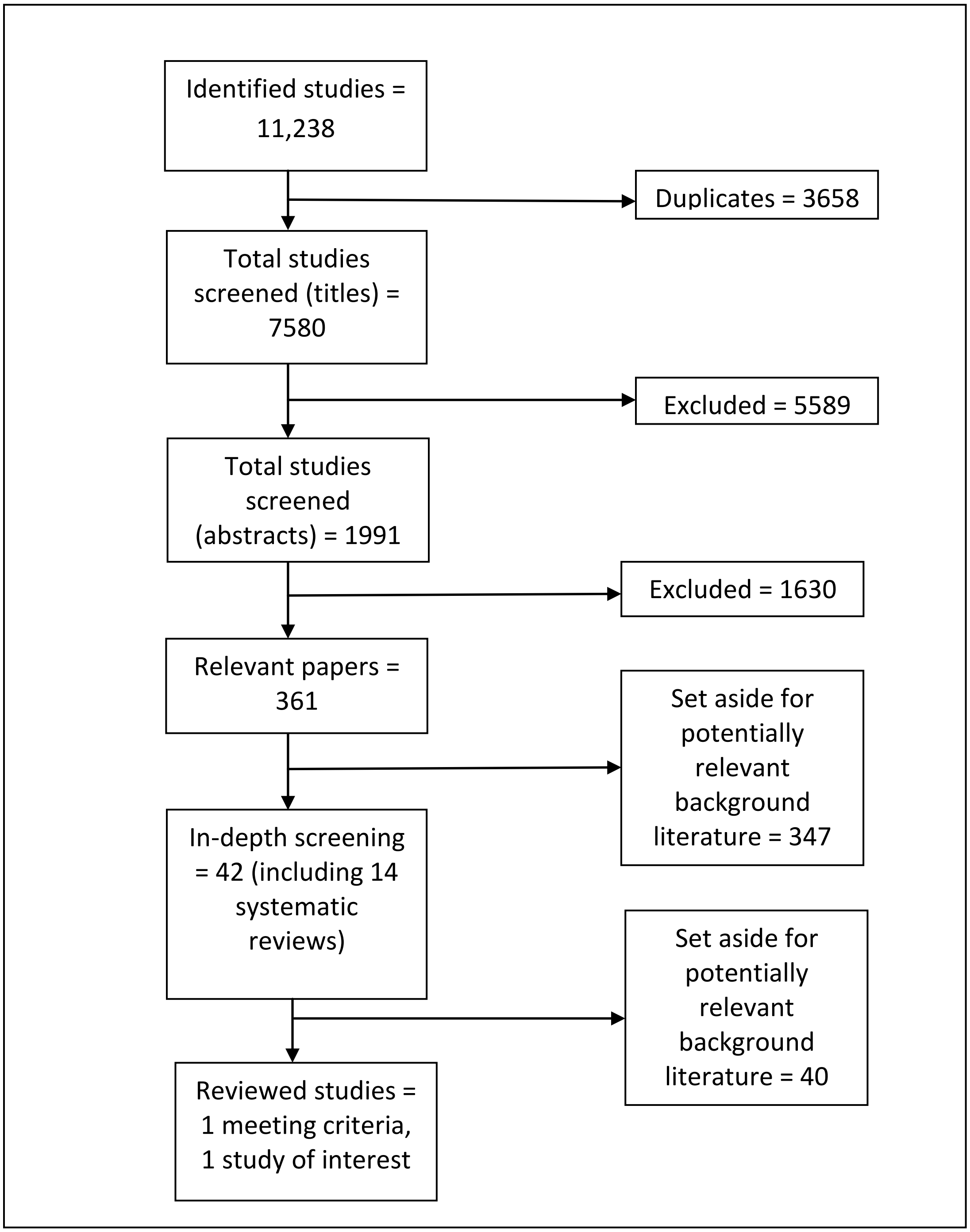


Table 3-2 Evidence of instrument properties

Name of measure	**Sub-scales**	**Psychometric development**	**Reliability**		**Validity**			**Sensitivity**
			Internal	**Test-retest**	**Content**	**Construct**	**Criterion**	
Web Trust Questionnaire (Harris et al., 2011, Harris et al., 2009)	1. Access to quality information 2. Personalisation 3. Perceived impartiality 4. Credible design	Principle component analysis, Factors determined using Scree plot (eigenvalue > 1), Varimax rotation 55.6% Variance Explained	✓	✗	Use of literature to inform items. No reporting of user testing.	✗	✗	Demonstrated difference between websites on two scales: Perceived impartiality and Credible design
Affective and cognitive response (Harris et al., 2011, Harris et al., 2009)	1. Coping 2. Worry	Principle component analysis, No further data	✓	✗	Use of literature to inform items. No reporting of user testing.	✗	✗	Unknown
(Modified) Patient Enablement Instrument (Yardley et al., 2010)	None specified	No evidence for modified instrument	✗	✗	No report of user testing	✗	✗	Evidence of sensitivity between groups
Satisfaction with web-delivered advice (Yardley et al., 2010)	None specified	None specified	✓	✗	No reporting of user testing	✗	✗	Evidence of sensitivity between groups

3.3.1 Paper 1 (Harris et al., 2009)

An experimental RCT on women's perceptions of risk associated with alcohol consumption and breast cancer met the specified inclusion (Harris et al., 2009). Women aged 18 to 46, were randomised to one of two websites containing identical high quality information but differing in visual design elements (known positive and negative 'trust' cues). The website containing positive trust cues incorporated a certified privacy seal (TRUSTe) and a certification label which may be attained through meeting a number of reliability and usefulness indicators (Health On the Net foundation: HON code). The second website contained negative trust cues including advertisements, pharmaceutical sponsorship and a donation button.

This study aimed to assess the influence of design based credibility cues on the acceptance or rejection of important health information. Amongst other measures (such as, eye-tracking and recall of information to assess engagement with website, response to the website message and influence of cues on subsequent alcohol consumption), the RCT included a post intervention questionnaire which is of particular interest to this review, the *Web Trust Scale*. A further outcome of interest which was included in this study was a set of anxiety response items which were directly related to the website. Items are discussed further below.

3.3.1.1 The Web Trust Scale

The Web Trust Scale (or the *Web Trust Questionnaire*) was originally constructed with 18 items assessing trust in online sources (Briggs et al., 2002). In a subsequent study (Sillence et al., 2007a), supplementary items were added after considering qualitative research assessing trust and readiness to act on advice provided by a website (Sillence et al., 2006, Sillence et al., 2004). The questionnaire is comprised of four factors, which were identified using exploratory factor analysis (Access to quality information, Personalisation, Perceived impartiality and Credible design) and more detailed descriptions of their psychometric properties are outlined in a more recent paper (Harris et al., 2011).

In the 2009 RCT identified, two factors in the Web Trust Questionnaire demonstrated that participants trusted the positive cue website more than those in the negative cue website group: credibility through impartiality ($F_{1,82} = 4.74$, $P = 0.03$) and credibility through design ($F_{1,82} = 4.92$, $P =0 .03$) (Harris et al., 2009). Web Trust scores were used in this instance to verify manipulation of the cues had been successful. No significant difference was found on the remaining two factors: Access to quality information or Personalisation.

The Web Trust Questionnaire does exhibit reasonable psychometric properties and has been developed using standard data reduction techniques. Items were derived following qualitative research which supports their content validity. Testing items among users would further improve content validity (for example through cognitive interviewing). Further psychometric properties such as test-retest reliability or criterion validity have not been reported.

3.3.1.2 Negative affective response scale (anxiety response items)

The 2009 paper also highlighted the use of a negative affective response scale. Negative affective responses (for example, “The material on this website made me feel...”) were rated using a seven-response scale from ‘not at all anxious’ to ‘extremely anxious’. Although limited evidence of measurement properties was available for these items in the 2009 paper, a subsequent paper on the development of the affective and cognitive response scale was published (Harris et al., 2011). This paper included the following variables: worried, at risk, anxious about the risks, optimistic, in control and able to cope. Items were prefaced with “Looking at the site made me feel...” to assess reactions to a website. Two factors were identified: Coping (three items, Cronbach’s alpha=0.87) and Worry (three items, Cronbach’s alpha=0.77). Two single item outcome measures were also created measuring trust (I trusted the site) and readiness to act upon the information provided on the website (I intend to act upon the advice).

3.3.2 Paper 2 (Yardley et al., 2010)

One further study which did not strictly fall within the review criteria, however, proved to be of interest. This study evaluated the effects of a website aimed at helping individuals manage minor respiratory symptoms without seeking medical advice and compared two websites using items specifically relating to the effects of the website (Yardley et al., 2010). Participants were randomised to one of two websites: the intervention website (Internet Doctor) which provided tailored advice using an automated digital triage system and the control website which contained static information constructed with the best existing advice for reducing use of services and improving patient confidence. Perceived ability to self-manage health, changes in use of health services and satisfaction with the website were assessed. Of the measures used in this study the modified *Patient Enablement Instrument* and the three item satisfaction scale were of interest.

3.3.2.1 The Patient Enablement Instrument

A modified version of the Patient Enablement Instrument, which was originally developed to measure confidence to self-manage an illness after a face-to-face consultation (Howie et al., 1998), was administered after participants used their allocated website. Modifications to the instrument changed wording from “As a result of your consultation,” to “Compared with before you read the Internet Doctor webpages,”. The original Patient Enablement Instrument consisted of six items measuring the respondent’s ability to understand and cope with their health following a consultation. This instrument was modified so that items related to the patient’s experience of using the control website and the ‘Internet Doctor’ website. From what can be ascertained from the literature, this modification has not been tested to ensure it is a valid and reliable indicator of enablement after an internet consultation. The internal consistency (Cronbach’s alpha) was not reported for the modified instrument in this study. The modified measure showed some sensitivity to differences between groups with Patient Enablement scores higher among the intervention group.

3.3.2.2 Satisfaction with web-delivered advice

A three item scale assessing satisfaction with web-delivered advice was included in the RCT. Items included: “The website gave me all the advice I needed”, “The website was helpful to me” and “I felt I could trust the website”. The three items demonstrated good internal consistency (Cronbach’s alpha= 0.89). Participants rated the intervention website more positively than the control website on the satisfaction scale. Other than the reporting of internal consistency, no further reports of the psychometric properties are made available.

3.4 Discussion

The purpose of this chapter was to ensure there were no existing measures which were suitable for assessing the impact of using a health-related website before embarking on the development of a new measure. This review was confined to measures which had been used in controlled trials. The results indicated that very few studies had employed the use of generic, standardised instruments with proven psychometric properties. Two studies of interest highlighted a total of four tools with which to measure various aspects of using a website.

The *Web Trust Questionnaire* was the most comprehensively tested tool among the four identified. This questionnaire measured four domains (Access to quality information, Personalization, Perceived impartiality and Credible design) and was used post intervention to test the success of trust manipulation cues. The questionnaire sub-scales have reasonably good psychometric properties and appear to be easy to complete.

The remaining three scales identified assess affective and cognitive response (Coping and Worry), patient enablement and website satisfaction. These measures have relatively little evidence in support of their psychometric properties.

None of the scales identified accounted for the range of potential effects of using health-related websites, particularly when considering the many features now available through Web 2.0. More specifically, they do not measure the impact of accessing different styles of information, for example, the potential effects which have been identified following exposure to online experiential information (Ziebland and Wyke, 2012).

One of the most important findings of the structured review documented in this chapter was that the measurement of the impact of using health-related websites is severely limited by the lack of a suitable measurement tool. Furthermore, those measures which have been identified have limited evidence of being based upon established psychometric principles and study findings should therefore be treated accordingly.

A further result of this review has been the identification of the best measure currently available for the research context. Although the *Web Trust Questionnaire* is not designed to measure the impact of using a health-related website, it is likely that a person scoring highly on the specified domains (Access to quality information, Personalisation, Perceived impartiality and Credible) will also experience positive effects of using the same website. As such, demonstrating moderate correlations of a new measure with all (or part) of the *Web Trust Questionnaire* may support its validity.

3.5 Conclusion

Developing a new instrument, based on established psychometric principles, will facilitate comparisons between various styles of information available online today. Such a tool will enable the impact of using health-related websites to be assessed in a standardised, valid and reliable manner. Developing a measure through the use of existing literature, qualitative analysis, expert-patient opinion and psychometric analyses can support evidence of aspects of validity and reliability. With this in mind, this thesis sought to develop a psychometrically sound measure, suitable for use in clinical trials, which could capture the impact of using health-related websites across a range of health groups.

<u>CHAPTER 4:</u> The development of the eHealth Impact Questionnaire item pool

4.1 Introduction

Health-related measurement scales, such as the proposed e-Health Impact Questionnaire (eHIQ), require a clear conceptual basis to inform item generation (McColl, 2005, Streiner and Norman, 2008). Item development has traditionally been informed through a ‘top down’ approach where developers referred to relevant literature or reused existing items when constructing an item pool (McColl, 2005). Consulting the patient in the development of a self-reported questionnaire, however, is important as the process may highlight issues not found in the literature or through consulting healthcare professionals. Terminology can also be interpreted differently among various populations and user involvement can ensure that items and response scales are meaningful to patient groups (FDA, 2009, EMA, 2004, Lohr, 2002).

As a result, emphasis is increasingly placed upon the importance of incorporating the patients’ perspective from the early stages of questionnaire design and it is widely acknowledged that the conceptual underpinnings of a measure must be explicit and empirically based (Streiner and Norman, 2008, Bowling, 2005, Kerr et al., 2010, FDA, 2009, Reeve et al., 2013, McColl, 2005). A mixed methods research design is therefore considered good practice in questionnaire development (Kerr et al., 2010, Ring et al., 2010, DeWalt et al., 2007). The steps taken during the development of the eHIQ item pool (Stage One of this body of research) reflect the considerations above and are documented throughout the chapter.

4.1.1 Objective

Stage One aimed to construct and pre-test a set of meaningful and psychologically relevant items relating to the impact of using health-related websites. To construct an item pool the following objectives were outlined:

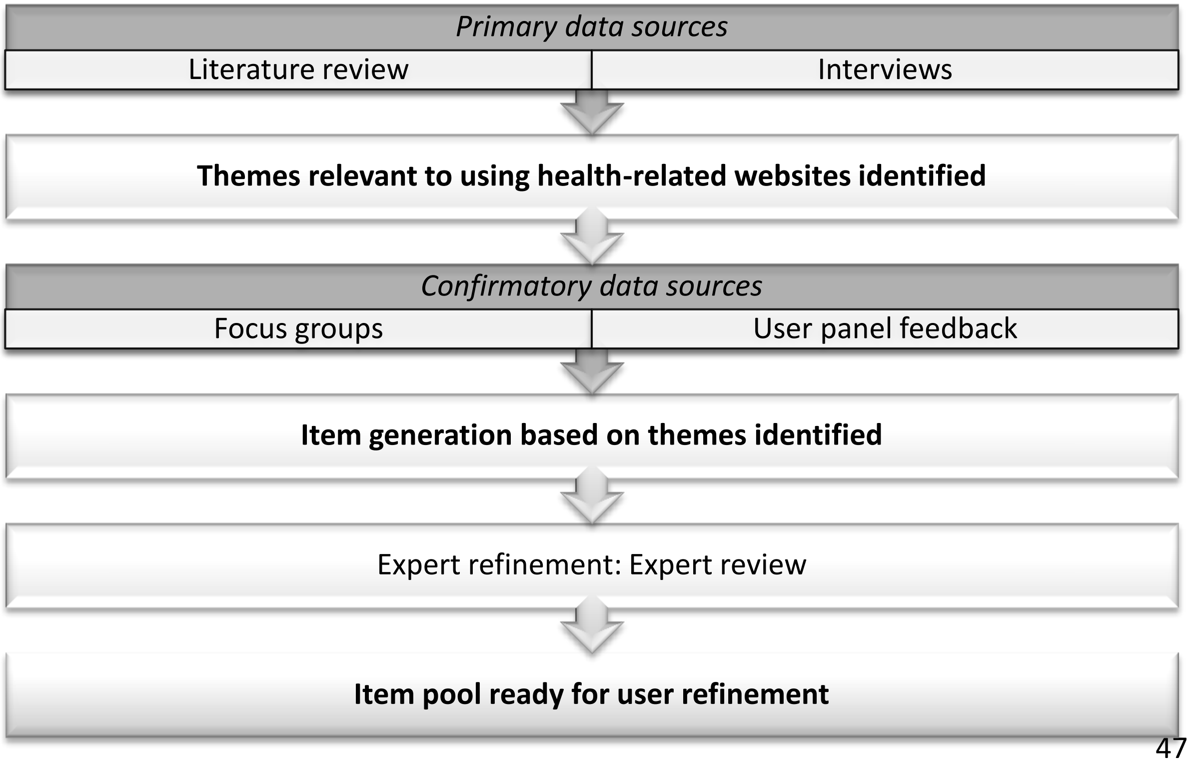
- To explore the main reasons why people use the internet in connection with their health and whether these reasons differ among health groups.
- To assess the impact of using various styles of online information.
- To select and re-cast 'statements' as questionnaire items which relate to the impact of using the internet (or websites) in relation to health. Statements needed to be applicable to using health-related websites incorporating one or more of the following: factual information, experiential information or online discussion forums.

4.2 Research design

A qualitative approach using existing data was considered appropriate. A theoretical framework using an existing review of relevant literature led the analysis (Ziebland and Wyke, 2012), while secondary qualitative analysis of narrative interviews relating to patients' and carers' experiences primarily informed items. Secondary data analysis is the re-use of data originally collected for another research purpose (Heaton, 2004). Access was granted to an existing archive of interviews belonging to the Health Experiences Research Group (HERG), University of Oxford, for the purposes of conducting secondary analysis. While the interviews were about people's experiences of illness, many contained accounts of how and why people use the internet in relation to their health and their reasons for seeking various forms of online health information and support. Using qualitative analysis allowed the incorporation of internet users' actual words as a basis for questionnaire items and helped enhance content validity. Content validity is important in questionnaire design as it establishes whether items adequately reflect the views of the population of interest (Brod et al., 2009).

Additional data sources were used to confirm the themes arising from the initial qualitative analysis. These confirmatory sources consisted of focus group transcripts relating to trust in websites and user panel comment forms relating to the effects of using the internet for health information. All data sources are outlined in detail in this chapter. Statements were selected to represent themes identified and recast as questionnaire items. A period of item refinement involving expert review followed. See Figure 4-1 for an overview of item generation steps.

Figure 4-1 Steps taken to generate statements and refine items



4.3 Data sources

Established qualitative data sets were used to conduct secondary data analysis. Heaton (2004) describes three forms of sharing data to conduct secondary data analysis; 1) Formal data sharing (data sets officially made available for sharing), 2) Informal data sharing (data acquired by request and private networks) and, 3) Personal or inside secondary analysis (the primary researchers reuse of own data sets to answer new research questions) (Heaton, 2004). This research employed all three modes of data sharing: formal data sharing through the use of the HERG interview archives, informal data sharing of focus group transcripts relating to trust and the internet which were provided by senior research colleagues (Nothumbria University), and personal or inside secondary analysis through the use of comment forms collected during a user panel discussion group where LK was present. All data sources are explained in more detail below.

4.3.1 Primary data sources: Literature review and subject interviews

4.3.1.1 Review of available literature

A realist review (Pawson, 2006) identifying existing research relating to the role of sharing online patients' experiences was used to inform and lead the qualitative research (Ziebland and Wyke, 2012). Realist reviews synthesise and evaluate relevant interventions in a structured manner through looking at how and why various mechanisms influence outcomes in particular contexts or settings (Pawson et al., 2005). Findings from the review highlighted existing research in the topic area and informed the study of the likely themes which could arise during the course of analysis. The findings of the review informed the coding guide and provided a theoretical framework to guide analysis.

4.3.1.2 Interviews: Formal data sharing (HERG archives)

Based in the University of Oxford, HERG have created an archive of narrative interview collections relating to patient and carer health experiences. In 2010, at the time of the study, the HERG database held interview collections for over 60 different health conditions. Interviews are recorded using digital video and/or audio recording equipment in the respondents home. Collections typically range from 40-50 transcripts aiming to achieve ‘maximum variation’ within the sample. In addition to academic and peer reviewed publications, some of which have focussed on how patients use the internet (Chapple et al., 2012, Hinton et al., 2010), the analysis of these indepth qualitative interview collections are published on two websites – www.healthtalkonline.org and www.youthhealthtalk.org. The Health Talk Online (HTO) websites have embedded hundreds of videoclips and audio extracts from the interviews. These clips illustrate the main themes from the interviews and are intended to be used as a resource for patients, their carers and health professionals in the relevant field (Ziebland and McPherson, 2006, DIPEx, 2014).

The HERG archives can also be made available for use in secondary analysis. Transcripts are requested and obtained by the secondary researcher through the archive administrator and a written agreement is signed by the secondary researcher. The HERG collections have been used for a number of other secondary analysis studies, some of which explore how people talk about using the internet (Lowe et al., 2009, Hunt et al., 2009).

4.3.1.3 Methodology employed by Health Experience Research Group researchers

HERG interviews are conducted using an open ended narrative structure followed by a semi-structured interview. The narrative interview allows the participant to relay their 'story' or experiences of illness whilst the following semi-structured section gives the interviewer a chance to clarify issues or probe to gain a deeper understanding of the participant's account. The interviewer may also introduce topics that did not arise naturally through the course of the persons' narrative (Ziebland and McPherson, 2006). Participants are usually asked about sources of health information or support, including the internet. The nature of the narrative interview was particularly useful for secondary analysis as a stringently structured interview style would have limited its use for research areas not closely related to the original research aims.

4.3.1.4 Transcription

HERG interviews are transcribed verbatim by a professional team who are asked to follow a systematic procedure of recording pauses, repetitions and murmurs (Ziebland and McPherson, 2006). The primary researchers then review the transcripts to add detailed non-verbal punctuation (for example, smiles, laughs, the length of pauses or places where the participant may be hesitant and other details that may be missed on the audio). Typically qualitative interviews which are carried out for the purposes of informing a health-related measure may be transcribed verbatim without detailed non-verbal punctuation (McColl, 2005). Less detail however may be appropriate for interviews which are analysed by the primary researcher who is likely to recall the context in which the interview took place. For the purposes of secondary research, the practice of documenting detailed punctuation can be particularly useful for the secondary researcher to gain an insight into the context of the respondents' answers. The majority of HERG interviews also offer the possibility to consult video footage of the interview, as well as audio.

4.3.2 Confirmatory data sources

4.3.2.1 Focus groups: Informal data sharing (The Northumbria discussion group transcripts)

Sixteen focus group transcripts were shared for this research by senior research colleagues from Northumbria University. The focus groups took place during 2004 and recruitment for the study was carried out through advertisements in local media. All participants used the internet at least once a week. Focus group participants were asked to spend 60 minutes searching for health information online, after which discussions centred on participants perceptions of trust and online health information (see (Sillence et al., 2007b) for detailed methodology). Due to their focus on assessing online health information, the transcripts were highly relevant to the research purpose.

4.3.2.2 Comment forms: Personal or inside secondary analysis (Service User Panel)

At the time of completing the literature review (Ziebland and Wyke, 2012) which guided this analysis, a user panel discussion was convened in November 2010. Thirty lay participants were recruited through contacting members on a list of people in the Oxfordshire Primary Care Trust (PCT) who had previously expressed an interest in taking part in research. Participants were selected with the help of a questionnaire to ensure a cross section of the community would be representated (for example, across gender, age, occupation, and ethnic group). The questionnaire also ensured the inclusion of participants who had used the internet in a variety of ways in relation to their health (for example, use of websites, forums, blogs) and a variety of health conditions were represented. The user panel were given an introductory presentation on various ways people use the internet in relation to their health and divided into groups to discuss how they thought people may be positively or negatively affected by online experiential information. A final comment form asked individuals to write down the most important ways in which online experiential information might affect people from both the perspective of a person with a condition and from the perspective of a carergiver. These comment forms were collated and made available for the purposes of this research.

4.3.2.3 Ethical approval

Ethical considerations were taken into account when using data sources. The HERG archive is collected with participant consent for use of their material for several purposes including secondary analysis by approved researchers. Transcripts were obtained through an agreement within the University of Oxford (see Appendix 4.1). Before transcripts are released, identifiable information is removed, consistent with the participant's preferences.

The focus group transcripts which were provided by researchers in Northumbria University were anonymised. At the time of conducting the research, participants gave permission in consent forms for their anonymised transcripts to be used for other research purposes.

4.4 Methodology

4.4.1 The HERG interviews

Before conducting secondary analysis on the HERG transcripts, it was critical to ensure the data being used were fit for purpose. This can be assessed through looking at the accessibility, quality and suitability of the data (Heaton, 2004). Access to this data set was approved through liaising with the archive administrator and satisfying existing ethical protocols. The quality of the archive was assured through the high number of peer reviewed articles based on the interview collections, the seniority and experience of the researchers and the competitive funding applications obtained to carry out the original studies. Suitability (or fit) of the data was assessed continually through preparation, sampling and analysis of the data.

4.4.1.1 Sample design

From the outset, a large number of interview collections were considered for inclusion in the sample. Secondary research involves the reuse of data not originally intended for the current research questions. As other studies using this methodology note (Dargentas and Roux, 2005, Lowe et al., 2009), participants may not have been asked about issues directly relevant to this research. Screening a large number of transcripts was therefore important to ensure discussion of the internet in relation to health arose within a substantial number. As the research aimed to inform a generic item pool, the sample required a diverse representation across condition groups to support validity.

The HERG archive contains interviews conducted across a range of conditions, demographic profiles and roles (i.e. carers and patients). Achieving a varied socio-demographic sample was therefore relatively easy as HERG interview collections usually reflect this through the initial efforts of the primary researcher. Purposively selected interview collections included interviews with carers, people with long term conditions and with those making health-related decisions (for example, those undergoing treatment or screening). Within these three broad health groupings, the conditions chosen aimed to reflect people experiencing symptoms relating to both physical and mental health (and their carers). Interview collections which were likely to incorporate a range of potential effects of using the internet, for example, those outlined in Ziebland and Wyke's review (2011), were also purposively sampled. The intention was to maximise the variability in the groups sampled who were making use of the internet. This decision was made based after reviewing the content presented on the Health Talk Online website and discussions with the primary researchers.

A series of steps were taken to assess the suitability of specific interview collections prior to their inclusion in formal analysis. These steps helped to determine which interview collections were likely to be rich data sources for the research questions and are outlined below:

1) Key findings of HERG interview studies are published online at www.healthtalk.org. These summarised online findings were explored to provide an insight into the themes or topics that the primary researcher had identified in each collection. Summary reports were compiled to identify modules which had content related to using the internet for health[1]. A condensed tabulated summary was subsequently produced by LK (see Appendix 4.2) to aid her familarisation with key topics raised and gauge the extent to which participants may have discussed using the internet in a collection of interviews.

2) The primary researcher of selected interview collections was consulted face-to-face or via email to ask specific questions about the interview context. It was helpful to gauge their general impressions of the interviews they conducted and their thoughts regarding their potential value in relation to addressing the research questions.

3) Following consultations on the suitability of an interview collection for the research purpose, primary researchers provided coding reports (generated at the time of their own research) which they believed would be particularly relevant to the topic. Coding reports took the form of electronic reports of sections of interview transcripts which related to a specific topic and were subsequently reviewed to inform the sample selection. Examples of coding reports were entitled ‘Web forums’, ‘Information’, ‘Support groups’ and ‘Information (carer)’.

Completing the steps outlined above aided familarisation with the data sets and informed the potential suitability of particular interview collections. Sampling and selection of

---

[1] Summary report compiled by LK and Fadhila Haeri Mazanderani

interview transcripts were not finalised at the outset, but, as is usual in qualitative analysis, remained continuous throughout the process. This occurs as the sample size should be directed by the research question and subsequently through data saturation (Ziebland and Wright, 1997, Pope et al., 2000). The vast majority of transcripts included in the analysis were selected after screening the entire interview collection for each selected condition. As analysis continued however, a small number of transcripts were included as they were thought to be rich data sources. These transcripts were either recommended by the primary researcher or sourced after reviewing the Health Talk Online website. These interview transcripts belonged to the following collections: Antenatal screening, Fetal abnormality and Menopause.

4.4.1.2 Choice of interview collections

Using the sampling framework outlined, a total of 198 interviews across ten interview collections were chosen to undergo analysis (see Table 4-1). All interviews were carried out during the years 2005 to 2010. Of the 198 interviews, 93 transcripts included relevant data (n=93, 46.97%) and these interviews were carried forward for further analysis.

Table 4-1 Stage One screened interview transcripts

Condition group	Year of original collection	Number of transcripts screened	Total included in analysis
Long term condition (younger people)	2007	30	15
Type 1 diabetes (younger people)	2006	37	19
Depression (younger people)	2009	32	7
Parkinson disease (carers)	2008	4	4
Motor Neurone Disease (carers)	2008	14	9
Dementia (carers)	2005	28	13
Antenatal screening	2005	5	4
Fetal abnormality	2006	12	12
Menopause	2010	10	6
Mental Health: Black and minority groups (carers)	2008	26	4
Total		198	93

4.4.2 Method of analysis

Interview transcripts were analysed using a modified version of the Framework method, an analytical approach developed by the UK based National Centre for Social Research (NatCen). Framework analysis allows a researcher to look at the data, and conduct analysis in a systematic and comprehensive manner (Ritchie and Spencer, 1994). Developed specifically for use in applied qualitative research, the Framework method is more structured than other types of qualitative analysis and involves five stages: 1) familiarisation with the data gathered; 2) identifying a thematic framework which allows emerging issues, concepts and themes to be listed; 3) indexing the transcripts according to the thematic framework; 4) charting the data through extracting and synthesising it in a manner which allows within-case and between-case comparison; and 5) mapping and interpretation of data (McColl, 2005, Pope et al., 2000). These stages are discussed in more detail below.

4.4.2.1 Data extraction and familiarisation

Full interview transcripts were screened to identify and extract text where participants discussed using websites containing factual health information or experiential information. To aid this process, all 198 transcripts were imported into the qualitative computer software package NVIVO (QSR_International, 2010) and a key word search[2] was performed using the 'query' function. All transcripts were read paying particular attention to sections of the interview which displayed a high number of key words. Interview extracts from 98 transcripts were coded in NVIVO and re-read to facilitate familiarisation with the data.

4.4.2.2 Identifying themes and indexing

The secondary analysis sought to gain a deeper understanding of existing ('anticipated') themes found in the literature whilst being mindful of any new ('emergent') concepts which arose. The theoretical framework provided by Ziebland and Wyke's literature review guided the identification of anticipated themes. Emergent concepts were also important as this analysis was intended to be broader than the impact of experiential information (the focus of the literature review), and aimed to also incorporate include the impact of scientific information. Codes (or indexes) were collated to form a coding book which was revised where necessary for each interview collection (see Appendix 4.3). This

---

[2] Key words: forum OR post OR Internet OR website OR computer OR experience OR feeling OR support OR share OR positive OR negative OR message OR sharing OR online OR net OR chat OR information OR talk OR same OR email

facilitated a systematic process of indexing within NVIVO. Coding reports produced using NVIVO were read thoroughly to identify overarching themes relevant to the use of the internet in relation to health. Analysis continued until no additional themes arose. Further interview collections from the HERG archive were reviewed for incidences of ‘new’ themes until it was evident that no additonal points arose.

4.4.2.3 Charting

Charting data is typically a process of lifting and synthesising data according to arising themes, aided by software such as EXCEL. This allows the researcher to make within and between case analyses. For the purposes of this research, charting the data included extracting the data *verbatim* to facilitate the use of participant’s own words when forming items. Additional comments were added to the document where necessary to keep quotes in context.

4.4.2.4 Interpretation and mapping of data

Expressing results in frequencies can be misleading in qualitative research which uses a diverse (not statistically representative) sample (Pope et al., 2000). Although care was taken to reflect some weight or frequency at which themes arose, the analytical process sought to identify and represent a broad spectrum of experiences of using the internet in relation to health. Therefore sole or rare experiences were included in the interests of exploring and gaining an in-depth understanding of the topic (Ziebland and McPherson, 2006). Themes were then examined to see if they could be applied across a range of conditions or were unique to the condition. Consistency tables (See Appendix 4.4) were used to track which themes arose in each interview collection. Where a theme was evident in all interview collections, it was included in the generic item pool. If a theme was only evident in one interview collection, it was eligible for inclusion after further consideration regarding its potential applicability across conditons.

This resulted in the identification of generic themes central to the impact of using health-related websites. Themes were cross-checked with domains identified in the literature (Ziebland and Wyke, 2012) to ensure that key concepts were represented. Final themes were checked for applicability across three condition groups and three different types of health-related websites to ensure their suitability for inclusion in a generic item pool (see Box 4.1).

Box 4.1 Suitability checklist for the generic item pool

Items should be suitable to:

- Those who had viewed a website containing experiential health information
- Those who had viewed a website containing scientific health information
- Those who had used a website containing a health forum
- People who had a long term condition
- Carers of people with a health-related condition
- Health promotion behaviours, screening and treatment decisions

4.4.3 Validating the findings

4.4.3.1 Confirmatory sources

Two sources of data were used to check the themes identified for inclusion in the item pool: 1) Focus group transcripts (n=16) from research carried out on trust and online health information in Northumbria University (Sillence et al., 2007b) and; 2) Comment forms (n=29) completed by members of an internet user panel consisting of lay persons using local primary health care services.

These data sources were used in order to verify findings from the literature review and secondary analysis research. Using additional sources for secondary analysis ensured themes were consistent with other research and helped to identify any additional themes or issues that did not surface through the initial work. By using more than one source of data, this served as a form of 'data triangulation' enhancing rigour within the research (Robson, 2002).

4.4.3.2 Northumbria discussion groups

Focus groups invite a number of participants to interact and explore views around the research topic. Facilitated by a group leader, focus groups use the interaction within the group to stimulate and direct discussion (Bowling, 2009). Interaction can encourage thoughts and exploration around the use of the internet among participants which may not have been otherwise discussed in a one-to-one interview setting. The focus group transcripts were analysed using a modified framework approach and, of particular interest to the analysis, was the presence of any new themes which arose.

4.4.3.3 User panel discussion group

The user panel comment forms asked people to list the potential advantages and disadvantages of using the internet for health information, with a particular focus on experiential information. Comments were collated in a single document in EXCEL to compare issues raised with the themes identified in the research to date.

4.5 Item generation

4.5.1 Representation of themes and identifying generic statements

Several steps were carried out to select statements which could potentially inform the item pool. Each theme identified through the analysis was represented by relevant statements (in the form of verbatim quotes) from the HERG archive transcripts. Statements were managed according to their theme in a tabulated summary (item tracking matrix) which identified the health condition from where it originated (FDA, 2009). This allowed each statement to be traced to its origin throughout the iterative process and the tracking of any changes, deletions and reasons for amendments. Statements which could be answered by people across health conditions (i.e. generic statements) were identified by systematically going through statements and identifying those that could be answered by the groups outlined in Box 4.1. Statements were collated in a matrix and recast as questionnaire items by LK, SZ and CJ.

4.5.2 Expert refinement

Items were reviewed by an advisory board consisting of one clinician and five academics with interests in the field of ehealth. Reviewers were asked to consider if items were answerable to those groups outlined in Box 4.1.

Reviewers were asked to flag items which they thought a person in the outlined criteria could not answer and to critique the items using guidance adapted from a questionnaire designer’s tool (Willis and Lessler, August, 1999) (See Appendix 4.5). In the interests of generating as many as possible items for the draft questionnaire additional or amended items were considered (Streiner and Norman, 2008, Olson, 2010).

4.6 Results

This section documents the findings of the analyses undertaken to inform and construct an item pool relating to the impact of using health-related websites. It begins with a detailed account of the results from the secondary data analysis by describing the population sampled and the themes identified. The themes identified were applicable across population groups sampled and were therefore considered suitable for inclusion in a generic item pool. The presence of further themes was investigated using the confirmatory data sources (focus group discussion transcripts and user panel feedback forms).

4.6.1.1 Participant characteristics (the HERG transcripts)

Ninety-nine participants, 28 (28.3%) men and 71 (71.7%) women, were included in the sample. All transcripts included some discussion of use of the internet in relation to a health issue. With the exception of four interviews conducted with couples and one interview with three young women, interviews were conducted on a one-to-one basis. Therefore, a total 93 transcripts were included in the analysis. Participants ranged from 15 to 80 years old and had a mean age of 35.0 years (SD 16.9). Carers accounted for (n=30) 30.3% of the participants interviewed whilst the remaining (n=69) 69.7% were interviewed about their own health. Of those who reported their ethnicity (n=75), 90.7% were white. Table 4-2 shows further detail.

Table 4-2 Stage One participant distribution by gender and condition

Condition group	Male	Female	Total
Long term condition (Younger people)	3	12	15
Diabetes (Younger people)	5	14	19
Depression (Younger people) *	3	6	9
Parkinson’s disease (Carers)*	2	4	6
Motor Neurone Disease (Carers)	3	6	9
Dementia (Carers)	5	8	13
Antenatal screening*	1	4	5
Fetal abnormality*	5	8	13
Menopause	0	6	6
Mental Health: Black and ethnic minority groups (Carers)	1	3	4
Total	28	71	99

* Interviews conducted with couples: Parkinson’s Disease (x2), Antenatal screening and Fetal abnormality. One interview conducted with three participants: Depression

4.6.1.2 Overview

Extracts in the HERG transcripts where participants discussed using health-related websites were identified. Health-related websites was the collective term used for websites which contain: factual health information, accounts or stories of people’s experiences of health, blogs about health or health discussion forums. Participants’ attitudes towards the internet as a source of information and support are described and the themes identified as relevant for inclusion in the item pool are outlined.

4.7 Using the internet as a source of health information and support

4.7.1.1 Use of and feelings towards online information

Some participants accessed health-related websites intermittently in relation to their health condition. Information needs for those with long term conditions, in particular, varied according to key health events. Common points where participants used the internet in relation to their health were: pre-diagnosis, diagnosis, progression of illness or at a time where they needed to make a health-related decision.

Several participants indicated they were aware of a range of online sources of information available to them (for example, conventional health websites, health discussion forums or blogs) and many used information in connection with advice given by a healthcare professional. Occasionally participants reported feeling uncomfortable if they had to rely on the internet as their only source of information and liked to be guided by professionals.

> *Well the internet's a Pandora's box isn't it? ... The descriptions [of symptoms] that other people put down vary hugely. The... association …of particular diseases with particular symptoms…this is why you need the professional...the internet's fine but you could go down so many rabbit holes. (Motor Neurone Disease - Carer) MND46*

Participants differed in the amount of information they wanted to access. Some participants described avoiding online sources which they thought may be upsetting or distressing. A woman recalled how she avoided information that would have a negative impact on her mental wellbeing:

> *I don't really tend to read much [online] that makes me feel kind of depressed or anxious...because I don't think that's very useful... (Menopause) MEN10*

Experiential information on health websites could be a powerful way of relating information to everyday life. Some participants wanted to source information on the practical implications that certain choices may have on their lives:

> *When we were trying to find out as much information about [the baby] …the most useful website we found to tell stories of families was, it was an American website called 'A Heartbreaking Choice'…some of the information was a bit over the top, but …it was the most useful website because it told us about families and how they were coping. (Fetal abnormality) EAP23*

Online support groups were beneficial for those who felt geographically isolated and for those who liked to be able to dip in and out of peer-to-peer interactions. For example, using the internet to interact with others as opposed to a face-to-face support group meant a person could control the amount of time spent interacting with a particular group. Whilst the internet gave the user a degree of control over interaction however, it could still be difficult for some who described being drawn to negative accounts:

> *... I go and read [online] and sort of zone in on probably the worst experiences so of course ... that plays on my mind a little bit. (Menopause) MEN19*

For some aspects of health, learning about other people's experiences was at least as informative for a number of participants as advice given by healthcare professionals:

> *...you just need somebody who…who's had it happen to them and come out on the other side …it's all very well speaking to doctors and nurses about it but they… don't have the condition...... a doctor or a nurse is going to give you all the, all the health jargon … (Type 1 diabetes) DYP36*

The interviews did not include many comments about pictures and videos on websites. However, where they did arise it was evident their use could make information more vivid and stay in a person's thoughts:

> *I looked on the Internet I think the day after we found out about the heart condition. I went to the Down's Heart Group. I suppose the, the first thing that struck me was –…the opening pages you see a little girl with Downs Syndrome. I suppose that picture has just always stuck in my head from the first time I saw it. (Antenatal screening) AN30*

4.7.1.2 Engagement with a website

Information on a health-related website needed to be relevant (for instance, to a person's circumstances, stage of illness, age group or, in some cases, gender) to engage the user. This seemed to be particularly important on health discussion forums where people interacted. For example, one man discussed how his wife found support through a website relating to fetal abnormalities, but few men posted on the site:

> *What has happened is [my wife] joined the ARC [Antenatal Results and Choices] Forum on the internet. And I joined… the men's forum as well, there's a men's forum, which is so under-used it's, it's not true - I think it's bloke thing [laughs] - we don't like to talk about it as much. (Fetal abnormality) EAP35*

In addition to the website appearing relevant, participants discussed the importance of finding trustworthy and reliable information online. Participants often used more than one website to source information and many were aware of the relatively unregulated nature of the internet:

> *Be careful where you collect the information from. The Internet's very good, but it's also very dangerous, with a lot of parent-made websites that have no medical background to them or proper research... It can be very frightening, very misleading and just untrue. (Antenatal screening) AN30*
>
> *...you don't know if it is reliable, it's [the internet] a very big place and there's all sorts of people writing all sorts of things. You have to be careful that where they're coming from is, what's the word…a genuine background. (Long term conditions) CI23*

Participants were sometimes cautious with health discussion forums and blogs as they could not verify a person's account of their health. Some participants described finding misleading or contradictory information on health forums as members conveyed conflicting views:

> *Just things like some people saying, "Oh, it's fine to have, you know, sugar, you know, diabetic chocolate and things" and then other ones saying, "Oh, no, you shouldn't have, have them". (Type 1 diabetes)DYP27*

4.8 Impact of using health-related websites

The literature review (Ziebland and Wyke, 2012) identified seven potential effects of seeing and sharing experiences online. This analysis supported the findings of the review and found five of the seven domains to be particularly applicable across all the conditions included in this analysis. These themes were thought to be particularly relevant to the impact of using to health-related websites containing scientific information and/or experiential information and are discussed in detail below. While the remaining two domains identified in the literature review (*Learning to tell your story* and *Visualising disease*) were less prevalent throughout the transcripts, issues relevant to each domain were spread among the five highlighted themes.

4.8.1 Information

Many participants used websites to learn about their health and increase their knowledge on specific aspects of a health condition. Several used the internet to instantly access information and typically consulted multiple websites.

> *...we became experts on trisomies and all sorts of genetic disorders...it's wonderful now with the internet because you just dial up you know 'genetics', or 'abnormalities' and you just go on this journey and find out absolutely everything there is to know.... (Fetal abnormality) EAP32*

Information was sourced online to improve knowledge about specific problems including legal or financial areas affected by a person's health. A carer, for example, sourced and collated information online about caring for a person with dementia:

> *I needed to reach out to find out more, at which point we joined the Alzheimer's Society. And that, the [online] Alzheimer Newsletter and various other things, … within about twelve months we'd built up quite a bank of information about the disease and about the process of caring and dealing with the various issues, legal, financial and so on. (Carer- Dementia) ALZ07*

Using health-related websites as a source of information at key health events (for example, a diagnosis) was particularly useful for some participants. One couple discussed using the internet when they learnt of complications with their pregnancy. Seeing examples of positive stories in this context could be quite compelling:

> *…we used the internet … to find as much information as we could on, on her condition. We found a very useful website that showed children who had gone through operations and had, had survived, which was very interesting for us. (Fetal abnormality) EAP23*

Although the internet was viewed as a valuable resource for instantly accessing information, participants sometimes reported difficulties in making sense of the mass of information they found online. Aside from knowing which information to trust, the quantity of information meant some participants found it difficult to keep up to date with information relating to their health:

> *I did put the time in but...a lot of it [information online] is very misleading… the research and the information and medicine is progressing and all you're doing is*

> *finding out more information constantly which will pre-date all that's happened previously. So it can be misleading because there's such a plethora of information. (Dementia carer) ALZ12*

Taking an active role in sourcing information helped some people cope when facing uncertainty for the future. Experiential information proved particularly useful for gaining future insights. This was applicable across groups such as those with long term conditions, carers and those facing health-related decisions. Although this activity could shed light on the possible progression of an illness, it could be an upsetting time for some participants:

> *...I did [look on the internet for information] when I was a bit younger but I think the information that I came across of sickle cell more or less tells you you're going to die at a young age. So it wasn't very encouraging for me to go back and look at more information...I was upset...So I just never went back to look at it again. (Long term condition) CI04*

Participants sometimes drew on experiences of others when making decisions which would affect their daily lives. Such decisions often related to issues that had important practical, ethical and emotional consequences that could not be fully explored solely with medical advice. When facing the decision of continuing a pregnancy with a fetal abnormality, for example, one woman described emailing a woman she met on the internet who had chosen to continue to full term birth to ask very specific questions:

> *I then sent her a long email asking [laughs] very specific questions about, "Okay, when you take him swimming you know what happens, do you take the prosthesis off, do you leave it in the locker and then go swimming?" [laughs] 'What happens when you get him out of the car - does it scratch the car?' ... it was just that I really wanted to get my head around what, what his life would be like, and what it might be like for me to be a parent of this child (Fetal abnormality) EAP37*

A number of participants discussed how websites were used for health-related advice and tips. Tips could be taken from conventional information, but were more frequently reported when using experiential information:

> *[Husband] was having trouble turning over in bed, you know, and somebody had written [online]... "buy...satin sheets" (Motor Neurone Disease- Carer) MND34*
>
> *...you know sometimes in people's own personal accounts there are real nuggets of wisdom there, or things, little things that have helped them. (Menopause) MEN10*

Some carers used the internet to benefit both their own health and that of the person for whom they cared for. In some cases, use of the internet allowed them to solve reoccurring problems which they had not been able to solve previously:

> *I found the incontinence thing a really big problem because I didn't know where to go... [it] was upsetting for me and was upsetting for [partner]...I had difficulty in getting a commode because I didn't know where to go, what to do or anything. I found out by going on the Internet... I don't know what I would have done without the Internet because I got most of my information there rather than anywhere else. (Dementia) ALZ08*

4.8.2 Feeling supported

The internet was a source of support for some participants with health concerns. Participants sometimes described feeling comforted and less anxious about symptoms when they discovered others had similar health experiences and symptoms to them:

> *... You start to think "What the hell's wrong with me, have I got some disease?"...Until I found the website and I read the forum and I thought "Jesus there's hundreds of women like me... this is quite normal"'. (Menopause) MEN12*

Some recalled thinking they were the only person with a particular health concern until they learnt of similar people on the internet:

> *My mum actually looked on the Internet to find out information about M.E.[Myalgic Encephalopathy or Chronic Fatigue Syndrome] and she wanted to find out if there were any kind of groups for just young people...And that was just the first time that I realised there was lots of other young people who had it. (Long term conditions) CI07*

One diabetic recalled thinking she was the only person who had a fear of injecting insulin. She described how he thought she would have benefited from knowing of other diabetics had similar experiences through online resources:

> *I think if you are suffering from like a fear of injecting I think it should be written more about because I've never seen anything about it. I just thought I was the only one, like, where I was just that scared of it I just thought oh it's just me, but there's loads of people out there, that are in the same position. (Type 1 diabetes)DYP16*

Participants who had difficult health experiences or symptoms found the internet helpful at times when they needed emotional support:

> *...something happens and you can just dab into the Internet and just read somebody else's [experience] ... it's instant reassurance. ... (Long term conditions) CI28*
>
> *… when it [ending a pregnancy] first happened I used to write, probably every other day emails to all these other people [connected to a website] and get replies back. But now I don't feel I have to write all the time. (Fetal abnormality) EAP06*

Instant support (during the day or night) was particularly important for those with unpredictable conditions such as mental health problems. In these instances participants could talk to others without leaving their house or during unsocial hours:

> *... it’s [the internet] very comforting because I don’t have to get dressed but I can still talk to people who have chronic fatigue or depression because of web support groups. (Depression)YPD10*
>
> *... I wasn’t sleeping and if I was up at three o’clock in the morning feeling the need to self-harm there was no-one there at all, and then suddenly this world on the internet has opened up people sort of in various different countries across the world or people who also weren’t sleeping, that were up at the same time, were able to just distract me or talk me through it ... they helped me through it, which was great. (Depression)YPD18*

Seeing how others cope with their health through online accounts or discussion forums encouraged some people to have hope. Participants described how health websites which portrayed a wide range of health experiences could highlight how illness can affect people in many different ways. A website which had a positive tone or included interactions between people who had optimistic views also helped some participants to foster a more hopeful outlook in relation to their health.

> *I was sort of like “Oh no” like, “What have I got to look forward to?” ...the net’s quite a useful thing, and you can talk to people, you don’t have speak to people face to face, … and you can hear different stories from people … that’s probably what got us, got us through … realise that you know, … it’s not a sentence. (Long term condition) CI19*
>
> *... we did come across a website …and the little mantra was, …that ALS [Amyotrophic Lateral Sclerosis] was not life-threatening, it was life-changing, lifestyle-changing. And we decided that that was a really good thing to stick to… so while our life perhaps hasn’t been the way we thought it might have been, it’s still been a good, it’s still been good. We’ve still done lots of things and you can still be positive about it, which is just really important. (Motor Neurone Disease- Carer) MND20*

Others however found it counterproductive to contrast their health with others. Some participants with depression, for example, found discussing sensitive issues on discussion forums upsetting:

> *...something that they [other forum members] say will just trigger you off…it is really sensitive, you have to be quite careful with the online groups... if you're feeling quite low and you go on line and you have a bit of a rant about something… you'll read someone else's rant, just general intro and it just makes you feel worse. (Depression) YPD17*

4.8.3 Relationships with others

In some cases websites which incorporated comment fields, ratings or discussion forums could reduce feelings of isolation and enable participants to feel understood. Those who indicated they had a very active online presence (for example, through interacting regularly in discussion forums or email support groups) described feeling a sense of community and comradeship with other users.

> *… I would provide people [in an online support group] with information and they'd come back with ideas and if we were going for consultations, I'd let people know …we were looking for… good vibrations coming through and it was a wonderful support for us...if you don't have that peer group, which, can give you that support it must be terribly difficult. (Motor Neurone Disease- Carer) MND46*

Some participants used discussion forums to offload concerns which they were unable to tell people in their everyday lives. Venting concerns to people who were not present in daily life allowed people to talk about concerns openly without offending people close to them and gave them some control over when they wished to discuss these concerns.

> *... the internet and forums as well were really helpful, you just post like a message, and like say your concerns, and then you go back maybe a day later, and there will be five or six people who, who've like says you know "I've been through this" ... if you feel like you can't talk to people around you, it's just that, a little bit more than like a resource that you didn't have. (Long term conditions) CI19*

Health websites allowed participants to access others who had uncommon conditions and specific co-morbidities. A young woman who had both diabetes and an eating disorder illustrated how it was difficult to relate to someone with *either* diabetes or an eating disorder. She felt understood by people who had both health concerns:

> *… But from like the perspective of having an eating disorder and diabetes, like that particular combination, it's been something that I've like really needed. Because like, like just talking to people who have an eating disorder, they don't understand the diabetes side of it. And like talking to people who just have diabetes, they don't understand the eating disorder side of it. [um] So I've actually like found several people online who have like the same combination. (Type 1 diabetes) DYP11*

Using the internet had the potential to make offline relationships easier through providing an online space where participants could vent their health concerns and not burden friends or family. One man however, described feeling hurt when he read a post which had been written by his partner on a discussion forum:

> *.. she'd shared a feeling... and my initial reaction was, "Why the bloody hell can you share that with that person you don't know, but you can't share it with me?"…. (Fetal abnormality) EAP35*

Carers in particular talked about how health-related websites could be beneficial to their relationship with the person they cared for. Carers discussed how tips and advice not only helped them practically as a carer, but helping to problem solve could also reduce stress. Learning ways in which they could make life easier for the person they cared for meant that the relationship between the carer and the person cared for was less strained.

Using health-related websites and reading other people's stories also assisted some participants in finding the right language to articulate to people in their everyday life how they felt and what they were going through. Strains in interpersonal relationships were sometimes reduced through communication:

> *I got a posting off of their forum …Somebody was asking advice for their husbands and partner and this lady posted this reply and it listed the thirty five symptoms of the menopause and how she thought the partner should help the woman. And I printed it off …and I said "[Partner], I've just found this do you want to have a read of it?" … and he said "It makes sense, it makes so much sense what this woman's written". (Menopause) MEN12*

It was evident that some overlap existed between 'Relationships with others' and the previous theme 'Feeling supported'. For example, feeling a connection with others who were having similar health experiences online may have reduced a sense of isolation experienced by an individual and made them feel like they were part of a supportive community. Therefore, this may have had added an overall sense of feeling supported. Similarly, improved relationships with others could be facilitated by improved communication due to using health-related websites as illustrated above. This theme therefore could also overlap with Ziebland and Wyke's (2011) domain termed 'Learning to tell your story'. Such 'overlaps' in themes were somewhat expected as they are similar in nature and were unlikely to be independent of each other.

4.8.4 Experiencing health services

Participants sometimes discussed consulting health-related websites when they were unsure whether and where to access appropriate health services. Information shared on websites helped people to decide if they should seek medical advice and helped participants identify serious concerns:

> *...you can ask them [people online] how they dealt with something or how they handled a situation, whether they've experienced something and [if] they think it's related to diabetes, and like kind of before like going to like asking your consultant… you can see like what other people's experiences have been... (Type 1 diabetes)DYP11*

Before consulting healthcare professionals, participants occasionally spoke of using the internet to ascertain important questions they should ask in order to get the most out of a consultation. Going to a consultation 'armed with information' helped patients to be articulate and become more involved in health decisions.

> *We wanted to go in armed with all the questions… the internet is a wonderful thing … it arms you with the questions that you are going to ask the consultant at your next appointment... (Fetal abnormality) EAP32*

One couple discussed how they used the internet to research what they could expect during future consultations having previously attended a scan where they felt unprepared:

> *We went onto the internet…[and] started doing lots and lots of research. So I guess by the second time we went back up for another cardiac scan, we were quite well armed with what we could expect next. (Antenatal screening) AN30*

Participants often combined information provided by health professionals and health-related websites to become aware of potential treatments they may need. This helped participants to take an active interest in managing their health and participate in decision making. This was particularly evident in long term conditions such as diabetes:

> *...after I'd first heard about [an insulin pump], I... looked on the Internet and spoke to people about the [insulin] pump...then I was bringing it up with my consultant at every appointment saying "This is going to be a good thing for me". (Type 1 diabetes)DYP13*

Several participants used the internet after consultations in order to check, reconfirm or corroborate advice given by healthcare professionals. This was particularly evident if a person was not satisfied with their treatment, side effects of treatment or the effect their health had on their lifestyle:

> *..they said I was going to have to raise the injections that I was doing …that was not going to happen if I wanted to carry on my golf. So me and my mum went onto some websites trying to find some sort of alternative to injections. And we found this website on the pumps we found out loads of research, went to the GP and… he didn't know anything about it...we found the [hospital]…in London... and got onto a pump... (Type 1 diabetes)DYP31*

4.8.5 Affecting behaviour

Participants, particularly those with long term conditions, described wanting to know the short and long term consequences of their lifestyle on health. Some implied that getting the right information could be motivational when managing their health.

> *[I would like to see online] ... complications for the person who hasn't looked after themselves and the health benefits for the person who has looked after themselves. I think that sort of thing would be fairly useful and fairly motivational for someone who has got the condition. (Type 1 diabetes) DYP36*

Information and support on health-related websites helped participants to take an active role in their healthcare. Many used the internet in combination with guidance from healthcare professionals to establish the optimum treatments for them and to alleviate side effects:

> *I did a bit of research on the on the internet mainly with the help of Menopause Matters and one of the ladies on there suggested I might try a different patch …I had a flare up…I went back to the doctors and she said … "You're probably allergic*

> *to oestrogen full stop". Again, a bit of research on the internet and somebody suggested that I try one that is a different source not a plant or a source from a plant… (Menopause) MEN12*

A couple found it helpful to monitor and manage the degenerative condition MND by using an online portal which allowed them to monitor their own health status and compare their quantified data with others who had the same condition.

> *...there's a new website… called PatientsLikeMe, which is brilliant…[all] you do is you put your profile up and in it you record the treatments, the symptoms, how long you've had it and, and all that sort of stuff … it gives you the opportunity to, to find out other patients that are in the same…onset as you are, i.e. bulbar, and the same symptoms using different drugs or different supplements and whatever. (Motor Neurone Disease- Carer) MND23*

It is conceivable that finding out what treatments people have received or what supplements people with the same condition have taken (for example, as in the case of the couple cited above) may spur people on to try out such remedies. Seeing the consequences of poor health management could also be influential for future health choices and encouraged participants to take steps to improve their health:

> *I think there could be more information about what happens if you don't look after yourself. I mean we all know like what happens if you do look after yourself. But there's not enough …I think there should be more scare stories, to scare people into looking after themselves…. I think some people need it to push them into looking after themselves. I think if people see what could happen to themselves if they don't, they'd be more inclined to look after themselves. (Type 1 diabetes)DYP35*

4.9 Confirmatory sources

Confirmatory data sources were reviewed in order to ensure that each theme identified had been fully explored and that no additional themes were evident. No further overarching themes were identified using the confirmatory sources; however, some themes were explored further.

4.9.1 Northumbria University focus groups

Although no new overarching themes were identified in the Northumbria University focus group transcripts, participants offered greater insight into the role of the internet in relation to the fourth theme (Experiencing health services).

Some participants acknowledged that consulting the internet could prevent unnecessary visits to the doctor by providing a source where people can look up information about a condition to see if it warrants medical attention. Some were concerned however that individuals might misunderstand information or be misled by an inaccurate source:

> *... I got a bit concerned when I was actually reading it [help forum] 'cos I thought these people should be asking these questions of their doctor not to some stranger who signs himself anonymous (Northumbria-Study3 Week2b)*

Concerns over who provided information (i.e. the trustworthiness of information) were common among participants. This applied both to factual and experiential information provided on websites.

4.9.2 User panel forms

User panel comment forms asked members to write down the most important ways in which online experiential information might affect people from both the perspective of a person with a condition and from the perspective of a caregiver. A summary table (see Table 4-3) was devised to document all points raised.

Similarly to the Northumbria University focus groups, participants in the user panel noted that consulting the internet can prevent unnecessary trips to the doctor and aid self-care. Concerns were expressed however that people may be in danger of relying on an incorrect interpretation of information on the internet rather than seeking medical advice.

The user panel comments covered many of the issues raised through the secondary data analysis of the HERG transcripts. Interestingly, they provided further depth to the third theme identified (Relationships with others). Members of the user panel were concerned that people could become heavily reliant on relationships formed through health discussion forums and may become isolated from the ‘real’ (or offline) world. Participants also highlighted tensions for carers when using discussion forums. Two carers were concerned about being vulnerable to bullying when joining patient discussion forums as existing members may think the group should be exclusively reserved for people with the condition in question.

Table 4-3 Themes arising from user panel discussion

Population	Positive	Negative
Patient	Information can be followed up with the GP and can help person to ask the right questions Source of instant information Can help people understand what is happening to them: ‘the forearmed is forewarned’ Help with self-care and assistance Useful for learning management techniques Helps to inform decisions Can allay fears May boost confidence Can provide reassurance Can give validation Can empower Reminds people they are not alone Can help person cope	Information may be inaccurate Can be influenced ‘by fruitcakes’ Can encourage self-diagnoses (hypochondria heaven) May isolate oneself from real world/ obsessive use Abuse of vulnerable people Information overload Commercial interests Becoming anxious/fearful False expectations Anti-medical profession/replacing medical advice Extreme postings (trolling) Everyone is different- does not take into account severity
Carer	Helps understanding/empathy Reduces isolation Provides useful information Provides insight without having to ask intrusive questions directly to patient Can provide encouragement Can help person cope	Advice contrary to medical profession Anxiety/fear Hard to distinguish fact from ‘waffle’ Carer may be bullied Breeching patients confidentiality Increase feeling of ignorance May create false ‘Hollywood view’- happy endings Misreading information Experiences are individual

4.10 Representation of themes

Following the completion of the secondary analysis and consulting the confirmatory data sources, results were cross checked with the review of relevant literature used to guide analysis (Ziebland and Wyke, 2012). Findings were consistent with those found in previous studies and broad themes were similar to those highlighted in the literature review. It is important to note that the identified themes were somewhat fluid and were not always distinct in their own right. Therefore, themes could occasionally overlap. For example, the domains *Feeling supported* and *Relationships with others* could overlap when considering how becoming connected with other people with similar health concerns may reduce isolation and therefore create a sense of feeling supported. Using the internet in relation to health services may have also affected the course of action a person chose to take in relation to treatment and therefore also affected behaviour. These observations were somewhat expected and some overlap was therefore inevitable in the items developed. Use of consistency tables (see Appendix 4.6) ensured all themes could be applied across the three broad health groups supporting their applicability for a generic questionnaire.

Aspects of the two further domains identified in the literature review (*Visualising Disease* and *Learning to Tell the Story*) were also present throughout the qualitative findings. For example, *Visualising Disease* was evident in one participant's (EAP23, p. 70) recollection of finding information and aspects of *Learning to tell the story* was represented in another participant's (MEN12, p. 77) account of improving relations or understanding with her partner. These domains were therefore represented within the items developed for the questionnaire. The secondary data analysis reported in this chapter also highlighted the variability in people’s attitudes towards using the internet as a resource in relation to their health. On occasions, participants indicated they had a ‘baseline’ or pre-conceived attitude towards using the internet for health information and sometimes indicated that their orientation to using the internet for health information and support may influence the style of information they sought. Therefore, a second item pool relating to general attitudes towards online health information was developed based on these observations. Overall, two item pools were constructed to represent 1) Overall attitudes to health-related websites and 2) Attitudes relating to the consequences of using a specific health-related website.

Statements (376), in the form of verbatim quotes, representing the identified themes for the item pool were drawn from HERG transcripts. Generic statements (149) which could be answered by people across health conditions were identified by LK. Statements were recast as questionnaire items and reduced to 67 items in an iterative process involving LK, SZ and CJ.

Fifteen further items relating to identified themes were constructed by the research team. Related questionnaires (relating to activation, internet health information and social capital (DuBenske et al., 2009, Dutta-Bergman, 2004, Hibbard et al., 2005, Hibbard et al., 2004, Harris et al., 2011, Williams, 2006)) were reviewed for relevant items relating to themes that were underrepresented. Fifteen items were adapted or constructed for inclusion in the item pool. See Table 4-4 for example items representing each theme.

Table 4-4 Example items according to theme

Theme	Example item
1) Information	I have learnt something new from this website.
2) Feeling supported	I feel I have a lot in common with other people using this website.
3) Relationships with others	I feel I have a sense of solidarity with other people using the website.
4) Experiencing Health Services	This website raises questions I might ask a doctor or nurse.
5) Affecting behaviour	This website encourages me to take steps that could be beneficial to my health.

4.10.1 Expert refinement

Minor amendments to the wording of the rubric and items were made to improve clarity following reviewers' comments. Amendments were made to two items following the advisory boards (one clinician and five academics with interests in the field of ehealth) concern that they were unsuitable for participants with low health literacy. Reviewers agreed that items covered the themes identified as relevant to the impact of exposure to health-related websites and that items were answerable across a range of health conditions and roles (i.e. by a patient or a carer). A total of 82 items therefore went forward to the next stage of testing.

Items were arranged into two item pools, the eHIQ-Part 1 and the eHIQ-Part 2. The eHIQ-Part 1 contained 45 items asking about the participant's general views of online health information and the eHIQ-Part 2 contained 37 items which asked about the participant's perceptions of the consequences of using a specific health-related website. See Appendix 4.7 for items belonging to each item pool.

4.11 Discussion

This chapter documents the steps taken to develop items to inform an item pool relating to the impact of health-related websites. Items constructed were checked for their applicability across long term conditions, carers and those using a website targeted at health behaviour change. Items were also applicable to websites featuring scientific information, experiential information and discussion forums.

Secondary analysis of interviews relating to experiences of health was guided by a recent literature review (Ziebland and Wyke, 2012) relating to the potential effects of sharing experiences online. Five themes were identified as relevant to the impact of using health-related websites containing scientific information and to websites containing experiential information: 1) Information, 2) Feeling supported, 3) Relationships with others 4) Experiencing Health Services and; 5) Affecting behaviour. Confirmatory data sources were used to triangulate the findings. Issues raised in the focus group transcripts and user panel forms provided more depth in relation to the negative aspects of using the internet in relation to health, for example, becoming isolated from the offline community or misdiagnosing symptoms. Using a range of sources to identify and confirm themes supported their inclusion in the proposed item pool. Consistency tables were used to demonstrate the presence of specific themes across conditions. This was essential to distinguish if a theme was in fact suitable for inclusion in a generic measure. After item selection, the item pool was then evaluated by experts in the area of e-health. Instructions, items, response options and the electronic format of the instrument were all considered acceptable by the expert panel.

The secondary data analysis also highlighted the various preferences people have towards using the internet as a resource in relation to their health. Some psychometric scales have been developed to assess aspects of a baseline acceptance to receiving information. Dispositions to seeking and avoiding health information have been assessed using the *Health Information Orientation scale* (HIOS) and have indicated that trust in information providers can differ according to a person's health beliefs and orientation towards health information (DuBenske et al., 2009, Dutta-Bergman, 2003).

Other tools which measure aspects of readiness for internet specific health information focus on an individual's capabilities for using online information. The *eHealth Readiness* scale, for example, assesses factors associated with the digital divide (including domains such as, 'Provision of the internet for health' and 'Personal capabilities for internet use') (Jones, 2013). The *eHealth Literacy Scale* also assesses capabilities or personal abilities when accessing information (Norman and Skinner, 2006).

Such measures can potentially inform referrals to eHealth resources; however, they are more suited for assessments between internet users and non-internet users. It was therefore considered useful to develop an independent item pool to inform a measure of general attitudes towards online health information (eHIQ-Part 1). Items were developed to assess attitudes towards online information among existing internet users with the view to offering the opportunity to assess whether specific population groups may be more or less inclined to use various forms or styles of information. Such a measure may also be useful as a baseline measure in a trial to ascertain if there were any differences in attitudes towards online health information from the onset.

The methods used to inform item generation reflect best practice guidelines for the initial stages of questionnaire development (Streiner and Norman, 2008, Reeve et al., 2013, FDA, 2009, Lohr, 2002, EMA, 2004). Gaining a rich and detailed understanding of the construct to be measured is ordinarily achieved through focused interviews with the relevant population. While this approach is manageable for condition specific measures, the proposed items for the eHIQ needed to be applicable across a range of health conditions and roles (i.e. patients and carers). The reuse of data has traditionally been more common in the quantitative sciences for example, through the use of survey data. More recently, however, secondary analysis of qualitative data has grown and there were many pros and cons to using this method of analysis which many authors have identfied (Heaton, 1998, Heaton, 2004, ESRC, 2012, Corti and Thompson, 2004).

Limited time and resources meant that it would not have been possible to conduct interviews across a diverse sample of people for this study. The opportunity to carry out analysis on a large number of interviews spanning a range of conditions was therefore particularly useful for this purpose as it allowed the inclusion of many more interviews than it would have been otherwise possible to collect in the timeframe.

Secondary analysis can however be restrictive in comparison to primary research where the interviewer can focus their questions on the issues of most interest to their own research agenda (Heaton, 2004). In some of the 93 sampled interviews the original researcher had not probed into the experiences of using health-related websites. Integrating multiple, purposively selected data sets was therefore vital in ensuring all potential themes were investigated thoroughly. Secondary data analysis has also been critiqued for lacking relevant contextual knowledge when the researcher was not involved in the primary research. To some extent, the availability of video and audio files of interviews helped to overcome this problem. Suitability of the data was also assessed through a number of steps before formal analysis commenced: 1) Thematic summaries and participant biographies prepared by the primary researchers were read, 2) Primary researchers were consulted to gauge the appropriateness of the data for the research purpose and; 3) Primary researchers' coding books of relevant themes from their initial analyses were consulted.

The interview collections chosen were carried out over a five year timeframe when the internet changed quite rapidly. The presence of data relating to using the internet for health information may therefore have been expected to be lower in older interview collections due factors relating website availability and accessibility. There was however no evidence to support this when comparing interview content to the years of interview. Nonetheless, as the internet environment changes very rapidly, the way in which people describe using the internet may have changed in the intervening period. The confirmatory user panel data used to check identified themes did not suggest the way in which participants discussed using the internet in the HERG transcripts had become outdated.

4.12 Conclusion

This chapter suggests that individuals who use the internet in relation to their health may be affected across the five key generic themes: 1) Information, 2) Feeling supported, 3) Relationships with others, 4) Experiencing Health Services and; 5) Affecting behaviour. Analysis also highlighted many individuals had pre-conceived health information preferences and therefore items were incorporated to reflect this.

All themes identified were applicable across a range of conditions and are therefore suitable for inclusion in the development of a generic item pool. Items relating to the identified themes have been incorporated into the item pool through including words used by the study population. Items were tested for acceptability among experts and were ready to enter further patient and carer cognitive testing.

CHAPTER 5: The pilot eHealth Impact Questionnaire: User testing and web survey design

5.1 Introduction

Following item development and expert review, the eHIQ items were arranged into an online questionnaire format. This chapter gives a detailed account of the iterative process used to design the online questionnaire and test item content among users (Stage Two). Face and content validity of questionnaire items can be supported through pre-testing an instrument among the relevant population. Face validity ensures the questionnaire superficially makes sense and can be easily understood (Jenkinson and McGee, 1997) while content validity ensures all aspects of the domain being measured are represented by the items (Bowling, 2005). In addition to using the conceptual framework to inform items, content validity is supported through evaluating patient understanding of items and determining whether these interpretations are relevant to the construct being measured (FDA, 2009, Bowling, 2005). Items which are not found to be directly related to the concept of interest are removed or amended.

The mode of administration and format are important elements of questionnaire design. Traditionally self-complete questionnaires are administered using pen and paper format. For the eHIQ however, a web survey was favoured over hard copy administration as the questionnaire was intended to be used in conjunction with an online intervention. By implication the respondent should have a basic internet competency and it would be convenient for the respondents to answer online after using the intervention website. From a developer’s viewpoint, advantages of a web based questionnaire included the ease at which amendments could be made, the potential to collect data swiftly and avoiding the need for manual data entry (Eysenbach and Wyatt, 2002).

Typically questionnaire design considerations include exploring the most appropriate question type, item wording, item order and format. In addition, web questionnaires differ in three ways: 1) they reside on the internet, 2) they are executed on a web server and, 3) respondents access it via a standard web browser (Baker et al., 2004). Questionnaires which are administered over the internet therefore require additional consideration regarding their usability and whether the design is appropriate for computer literacy levels among the likely respondents (Baker et al., 2004, Dillman and Bowker, 2001). Computer competencies are difficult to forecast. However improvements can be made to ease completion and reduce non-response. Pre-testing the online format of the new questionnaire was therefore considered appropriate so that completion rates were not negatively influenced due to technical difficulties (Anderson and Klemm, 2008).

5.1.1 Objectives

Stage Two aimed to design and pre-test the online eHIQ with an appropriate patient and carer sample. The objectives were as follows:

- To ensure the instructions and items were easy to understand
- To establish face and content validity of the items
- To ensure the appropriateness of the response options
- To ensure the suitability of the online questionnaire format

5.2 Research design

The eHIQ item pool was arranged into an online questionnaire format using appropriate survey software. Cognitive debriefing interviews were used to explore respondents understanding of items (Willis, 2005b, Willis, 1999). Cognitive interviewing served to support the content validity of each item by assessing participant understanding across a large range of population groups (FDA, 2009). Face-to-face interviews also enabled the format of the web questionnaire to be examined.

5.3 The questionnaire design

Web survey software can offer a variety of features which need to be considered. These include formats, screen layout, colour, the use of filter questions and technical performance (Baker et al., 2004). Web surveys are also completed in a wide variety of uncontrolled environments, for example, unknown PC configuration, internet connection and browser software. Simplicity in design can therefore be preferable to achieve the most standardised questionnaire (Baker et al., 2004, Sue and Ritter, 2012).

After reviewing a range of products, Bristol Online Survey (BOS) software was chosen to construct the eHIQ. BOS software was considered easy to use, had reasonably flexible design features and had a professional appearing interface. A further advantage was that the software package did not place restrictions on the number of surveys collected, the number of items asked or the number of response options offered. The software was available at no cost to the study due to a license held by the University of Oxford.

The two item pools identified in Stage One were formatted in BOS software. Items were arranged into two questionnaire parts so that they would be easy to complete and would follow a logical order. A Likert type response scale of ‘Strongly Agree’ to ‘Strongly disagree’ was provided as a response options to the items. Likert scales are commonly used in attitudinal research and have a ranking order (Jamieson, 2004). Following completion of eHIQ Part 1 and Part 2, participants were asked to complete some brief demographic questions.

Various layout considerations were taken into account when designing the online questionnaire. The first page of the questionnaire contained the title of the study, a brief description of the aims of the study and eligibility criteria. The University of Oxford logo was inserted on the top right corner of every page to enhance appearance and credibility. Consistency and readability of the questionnaire were prioritised throughout the survey development and attention was drawn to specific words using font size and bold font (Sue and Ritter, 2012). As forced response options may have resulted in non-completion, participants were permitted to continue with the questionnaire despite not having checked all boxes (Sue and Ritter, 2012). A page number was inserted on each page to indicate survey progression and items (10-12 per page) were displayed in a grid (or matrix) format (see Appendix 5.1). Reviews and amendments to the questionnaire format were continuous throughout the cognitive interviewing process.

5.4 Methodology

5.4.1 Participants

Men and women (aged 18 years or over) who had a health condition or cared for someone with a health condition were eligible to take part in the study.

5.4.2 Procedure

5.4.2.1 Recruitment

Participants were recruited through contacting members on a list of people which was held by the Programme coordinator of the grant funding the research[3]. The email list had been collated through contacting patients in the Oxfordshire Primary Care Trust who had indicated they would be interested in taking part in research. Available demographic information for list members was used to ensure a range of conditions and ages were contacted. Potential participants were asked to contact LK to arrange a suitable time to be interviewed. Interviews took place in the then Department of Public Health, University of Oxford[4] or in the participant's own home. In addition to contacting people via an email list, participants were approached to take part in the study if they were known to the researcher as someone who may be eligible to take part in the research. This took place through word of mouth and through existing participants suggesting friends who may be interested in taking part in the study (i.e. snowball sampling).

---

[3] Funded by National Institute for Health Research, Programme Grants for Applied Research, RP-PG-0608-10147. Entitled: Examining the role of patients' experiences as a resource for choice and decision-making in health care.

[4] The Department of Public Health has since been restructured and renamed The Nuffield Department of Population Health

5.4.2.2 Ethical approval

Ethical approval was obtained through the University of Oxford Research Ethics Committee (Reference number: MSD/IDREC/C1/2011/98).

5.4.2.3 Recruitment materials and participant consent

Potential participants were provided with a Participant Information Sheet (see Appendix 5.2) explaining the background to the research and what they would be asked to do if they took part in the study. Before taking part, participants were given the opportunity to ask questions about the research and asked to complete and sign a consent form (See Appendix 5.3).

5.4.2.4 Cognitive interviews

At the time of interviewing, participants were reminded that the objective of the current stage of the study was to scrutinise the questionnaire. Interviews lasted approximately one hour in length, were audio recorded and notes of participant comments were taken throughout. Participants were given a payment of £20 for taking part and any expenses were reimbursed.

Participants were asked to spend 10-15 minutes on a pre-selected health-related website that was relevant to their primary reported health condition (see Table 5-1 for each condition and respective website). A variety of different websites were used including government websites (for example, NHS Choices), charity websites (for example, Health Talk Online) and commercial websites (for example, BootsWebMD). Websites were chosen to ensure the eHIQ items were tested both on experiential information and scientific information. Websites were also included to incorporate features such as discussion boards, video clips and rating scales.

The ‘verbal probing’ style of cognitive interviewing was used to allow respondents an opportunity to give uninterrupted answers to items, followed by a focused interview. This method of pre-testing items with participants is consistent with current guidelines relating to questionnaire design (FDA, 2009, Lohr, 2002). During the focused interview, the participant was reminded of how they had responded to a particular item and asked various questions in order to gain a deeper understanding of the reasoning behind their answer (See Appendix 5.4 for Interview topic guide). For example, in many cases probes such as: “How did you arrive at that answer?” were used (Willis, 2005a, Wilson, 2005). This method of interviewing not only queried the understanding of an item but also took into account the interpretation of the instructions and the response options provided (McColl, 2005).

5.4.2.5 Sample design

Participants were purposely selected to reflect a spectrum of health conditions and carers to test the questionnaire across conditions (McColl, 2005). Carrying out cognitive interviews served to identify any problems a participant had with completing an item and was not intended to represent the frequency or extent of a particular problem in the general population (Beatty and Willis, 2007). A sample size of 20-25 participants over a series of interview rounds was therefore likely to be sufficient for the purposes of this study (McColl, 2005).

5.4.3 Analysis

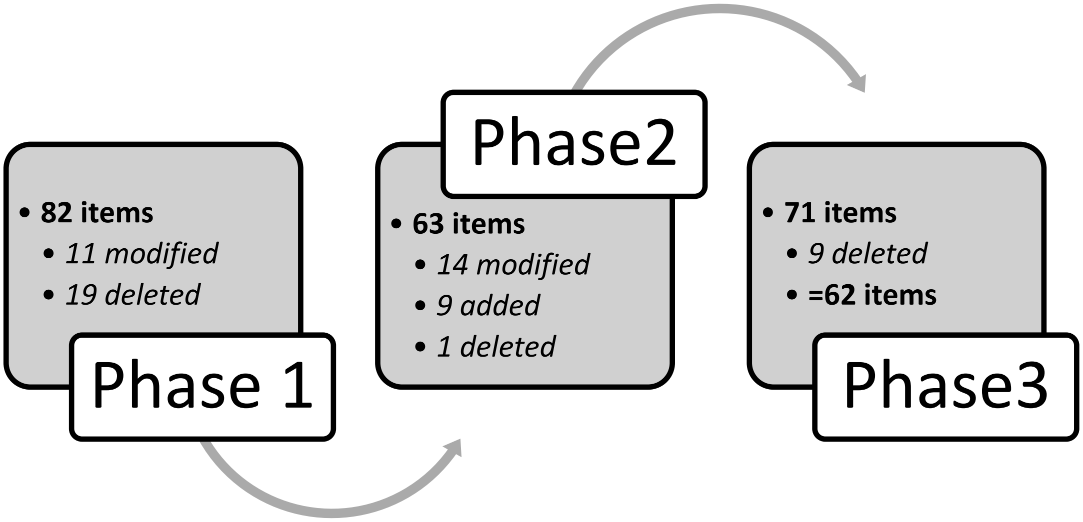
Individual participant comments for each item were documented and combined with interview field notes in a Microsoft Word template listing all items. All comments were subsequently combined in one document allowing comparison between participants. Consistency of interpretation for each item was analysed paying particular attention to any anomalies across condition groups. Problems with specific items or wording were highlighted and attention was given to the response options to ensure the scaling was appropriate. This assessed whether the response categories were suitable for the item stem and that the midpoint response category was interpreted as the point between agree and disagree (i.e. as opposed to being interpreted as a 'not applicable' response) (Streiner and Norman, 2008).

Interviews and analysis were carried out in phases which aimed to have 5-15 participants in each so that revisions could be made to the questionnaire where appropriate and subsequently retested (Beatty and Willis, 2007). Interviews were conducted until no new problems were identified. Throughout the revision process care was taken to retain a sufficient number of items to represent the themes identified through the initial qualitative research. Finally, participants were asked to comment more generally to investigate whether they felt any important issues were not covered adequately.

5.5 Results

Stage Two aimed to pre-test the online pilot questionnaire among a range of health groups. Interviews were conducted with 21 participants over three phases. This section documents how the 82 items were reduced and amended to 62 items. All remaining items were considered easy to understand across population groups. See Figure 5-1 for an overview of the modifications to items during cognitive interviewing. The reasons for item modifications in each phase are discussed in detail throughout this section and a list of the items deleted is given in Appendix 5.5.

Figure 5-1 Stage Two Modifications to eHIQ items



5.5.1 Characteristics

Participants (n=21) were six men and 15 women with a mean age of 45 years (SD 16.2). Five participants were carers and 16 participants had a specific health condition (see Table 5-1).

The first phase of interviewing consisted of three men and eight women with an average age of 50.8 years. Participants, three of whom were carers, represented a range of conditions and had all completed secondary level education. Six participants were employed, one was unemployed and four were retired.

The second phase of interviewing took place with four women and one man with an average age of 37.2 years. Of the five participants taking part in Phase 2, two participants were carers. Four participants were in employed and one was a student. All participants in Phase 2 had completed third level education (i.e. Degree level).

The final phase of interviewing took place with two men and three women who had an average age of 41.2 years. Two participants had retired due to ill health, two were in employment and one was unemployed. All participants had completed secondary level education. No carers volunteered to take part in the final phase of interviewing, however, this was not thought to compromise the results as the previous phases had not suggested differences in carers' interpretation and response to items when compared to people in other health groups.

Table 5-1 Stage Two Participant and website characteristics

Sample	Age	Condition	Website viewed	Website domain
Phase 1 (n=11)				
Male	58	High blood pressure	Health talk online (Hypertension)	Charity
	68	Stroke	Health talk online (Stroke)	Charity
	18	Carer (Brain haemorrhage)	WebMD (Brain haemorrhage)	Commercial
Female	62	Arthritis	NHS Choices (Arthritis)	Government
	48	Depression	NHS Choices (Depression)	Government
	66	Policythemia	Webmedicine.net (Policythemia)	Commercial
	47	Carer (Mental illness)	Carers forum (Mental Health)	Volunteer website
	50	Urinary tract infections	NHS Choices (Urinary tract infections)	Government
	67	Carer (Dementia)	Health talk online (Alzheimer's-Carers)	Charity
	19	Daily fainting	NHS Choices (Fainting)	Government
	56	Breast cancer	Macmillian (Radiotherapy)	Charity
Phase 2 (n=5)				
Female	28	Smoking cessation	Smokefree NHS	Government
	51	Carer (Dementia)	Carers forum (Dementia)	Charity
	27	Coeliac disease	NHS Choices (Coeliac)	Government
	29	Carer (Diabetes)	Iddt.org/carers-corner	Charity
Male	51	Smoking cessation	Smokefree NHS	Government
Phase 3 (n=5)				
Male	52	Multiple Sclerosis	NHS Choices (Multiple Sclerosis)	Government
	30	Asthma	NHS Choices (Asthma)	Government
Female	59	Fibromyalgia	Health Talk Online (Chronic Pain)	Charity
	27	Asthma	NHS Choices (Asthma)	Government
	38	Multiple Sclerosis	NHS Choices (Multiple Sclerosis)	Government

5.5.2 Phase 1

Phase 1 interviews began by asking participants to browse a relevant health-related website for 10-15 minutes. Participants were then asked to complete the full online questionnaire (eHIQ-Part 1 and 2 in succession followed by demographic information). Participants were given a five point Likert response scale as follows: Strongly agree-Agree-Uncertain-Disagree- Strongly disagree. Analysis of interviews highlighted a number of potential problems which needed to be addressed. Subsequent amendments are outlined below.

5.5.2.1 Questionnaire structure

Participant comments suggested respondents found the sequence of browsing the health-related website and subsequently completing the full questionnaire confusing. Participants felt they were influenced by the website they were asked to browse when completing items in eHIQ-Part 1 (General attitudes towards online health information). Bias may therefore have been introduced to responses to eHIQ-Part 1. An amendment was made to the sequence of the questionnaire so that participants answered eHIQ-Part 1 first and were then asked to browse the health website. After browsing the specified website, participants were asked to complete the eHIQ-Part 2.

5.5.2.2 Amendments to items

Eight items within the eHIQ-Part 1 were revised to reduce ambiguity and to ensure that all items clearly related to health information on the internet (see Table 5-2, Phase 1). Three items within the eHIQ-Part 2 were also modified to improve clarity.

Seventeen items within eHIQ-Part 1 were deleted as participant interpretations did not confirm what these items were thought to measure. For example, two items were thought to examine 'attitudes towards general health information', not specifically attitudes to online health information and therefore compromised content validity (see Table 5-2, Phase 1). Two items were found to repeat content in other items (i.e. duplications) and were therefore removed.

5.5.2.3 Response options

The response set ranging from 'Strongly agree' to 'Strongly disagree' was found to be acceptable by all participants. During the course of the interviews however, further probing highlighted that some participants did find the midpoint response category 'Uncertain' to be slightly ambiguous. In these cases they did not always associate this category as referring to the midpoint of the five point scale. It was important to establish a clear midpoint category so that participants would refrain from misusing this category by interpreting it as 'Don't know', 'Not applicable' or using it where they did not understand the question (Streiner and Norman, 2008). The midpoint response category was amended to 'Neutral' with the hope of aiding clarity.

Participants were also asked to comment on why they had selected a particular response category to a particular item. This was helpful to establishing the scoring direction for each item.

Table 5-2 Stage Two Examples of item modifications, deletions and supplements

Problem	Original item	Action taken	Amended (or remaining) item
Phase 1 Clarity	There are people out there who understand what I am going through.	Reword item to ensure items are clearly referring to people on the internet.	There are people who use the internet who understand what I am going through.
	I think that others have had similar worries to me when making health related decisions.		The internet is a good way of finding other people who are facing health related decisions I might need to make.
Content validity	I prefer to receive extensive information about my health.	Delete	N/a
	I prefer to receive little information about my health.		N/a
Phase 2 Clarity	I would use the internet if I needed to make a decision about my health.	Examples of decisions provided to improve clarity	I would use the internet if I needed help to make a decision about my health (For example, whether I should seek medical advice or seek alternative therapies).
Missing	Underrepresentation of theme *Experiencing health services.*	Addition of item	This website makes me want to play a more active role in my healthcare.
	Underrepresentation of theme *Relationships with others*.		This website gives me the confidence to discuss my health with the people in my life (For example, my family or people at work).
Phase 3 Duplicate items	It is difficult to know what health information to trust on the internet. Similar to: I can work out what health information to trust on the internet.	Delete item duplicate item	It is difficult to know what health information to trust on the internet.
	After looking at this website, I worry more about things that I cannot control in the future. Similar to: This website makes me worry about things that are out of my control.		After looking at this website, I worry more about things that I cannot control in the future.

5.5.3 Phase 2

On entering Phase 2, the eHIQ contained 63 items in total: 26 items within eHIQ-Part 1 and 37 items within eHIQ-Part 2. Phase 2 aimed to test the restructured questionnaire and the amendments made in Phase 1 with a new sample of participants. Participants were now asked to complete eHIQ-Part 1, browse the specified external health-related website and then complete eHIQ-Part 2.

5.5.3.1 Rubric and instructions

Following the completion of five interviews, it was apparent that participants were having some difficulty navigating from the web survey to the external health-related website and back again to complete eHIQ-Part 2 items after browsing the specified website. Instructions given to participants immediately after completing the eHIQ-Part 1 were therefore made more explicit to help participants. This aimed to help participants to remember they should return to complete the remaining items within the questionnaire once they had finished browsing the specified website. Instructions were written in a series of steps with important words highlighted in bold.

5.5.3.2 Amendments to items

Six items within eHIQ-Part 1 were modified to reduce ambiguity. Amendments included improving wording or adding examples. Adding examples to some items was useful where participants had difficulty thinking about behaviours which were not immediately apparent such as decision-making (see Table 5-2, Phase 2). Eight items were also modified in the eHIQ-Part 2 to improve clarity to the response direction. One duplicate item was deleted.

5.5.3.3 Additional items

Following the various amendments made, three items were added to the eHIQ-Part 1 and six items were added to the eHIQ-Part 2 to ensure all the themes identified in the qualitative analysis were still adequately represented. Examples of items added to represent previously underrepresented themes are shown in Table 5-2, Phase 2.

5.5.3.4 Response options

Whilst participants were able to recognise the term 'neutral' as a midpoint category, some commented that the more conventional use of 'Neither agree nor disagree' would be preferable. 'Neither agree nor disagree' was therefore tested for acceptability during the next phase.

5.5.4 Phase 3

On entering Phase 3, the eHIQ contained 71 items in total: 28 items within the eHIQ-Part 1 and 43 iems within the eHIQ-Part 2. Phase 3 tested the all amendments carried out in Phase 2.

5.5.4.1 Duplicate items

Participant feedback suggested nine items where repetitive and had very similar interpretations to other items in the questionnaire (see Table 5-2, Phase 3). They were therefore deleted.

As no amendments were made to retained items or the overall eHIQ structure during Phase 3, no further cognitive interviews were carried out. Analysis of remaining items showed a consensus of interpretation and understanding.

5.5.5 The final item pool

Following patient and carer refinement, the final number of items remaining totalled 62. Items formed two independent item pools. The first item pool (eHIQ-Part 1) contained 23 items asking participants about their general attitudes towards online health information whilst the second item pool (eHIQ-Part 2) contained 39 items asking participants about the impact of a specific website they had used (see Appendix 5.6 for full list of items and layout).

5.6 Discussion

Stage Two of this body of research aimed to design and pre-test the eHIQ pilot questionnaire. Three phases of cognitive interviewing were carried out to ensure: the questionnaire instructions were easy to understand, participants found the items relevant and acceptable to answer, the response options were appropriate to the item stem and the electronic format was appropriate for use among a range of participants. Twenty-nine items were deleted and nine items were added in total, leaving 62 items to enter psychometric testing. Emphasis was placed upon retaining a sufficient number of items to represent each of the five themes identified in the qualitative analysis (i.e. 1) Information, 2) Feeling supported, 3) Relationships with others, 4) Experiencing Health Services and; 5) Affecting behaviour).

The pilot questionnaire consisted of two independent questionnaire parts. The first questionnaire (eHIQ-Part 1) contained 23 items and asked respondents about their general attitudes towards online health information while the second questionnaire (eHIQ-Part 2) contained 39 items asking the respondent about the perceived impact of using a specific health-related website. All items had a five point response scale (Strongly agree- Strongly disagree) with ‘Neither agree nor disagree’ selected as the midpoint category.

Carrying out cognitive interviews with a range of health conditions and carers marked an important step in establishing face and content validity. Changes were iteratively made to the questionnaire in order to improve the acceptability of the items and to ensure all themes identified as relevant to the impact of online health information were represented. This meant that, where applicable, any changes to the wording of questionnaire items were tested again for their acceptability and relevance. If this re-testing indicated an item was still not interpreted in the intended way, the process of amending and re-testing was repeated until a satisfactory result was achieved.

The ‘verbal probing’ method of cognitive interviewing was appropriate for testing the questionnaire among users. Other methods of cognitive interviewing such as the ‘think aloud’ method can affect the participant’s response to items. The act of speaking and describing your thought process aloud while completing a questionnaire may alter thought processes and time spent on each item. Verbal probing allowed the participant to complete the questionnaire as they would do in real life and allowed observations to be made in relation to the ease of the participants' navigation and questionnaire completion. Carefully selected probes focussed the interview and kept the participant on the right course of inquiry. Verbal probing can however be open to interviewer bias as they may direct the interview to highlight what they view are the most important issues (Beatty and Willis, 2007). Therefore, while it was important to follow up important points made by participants regarding ambiguities, care was taken to allow the participant enough time to raise any points they felt had not been covered during the interview.

It has been suggested that cognitive interviewing may not result in *measurably better* items, however, they may provide insights into finding the optimal way of asking a question (Beatty and Willis, 2007). The cognitive interviews were therefore a valuable step in laying the foundations to optimal questionnaire design.

The Likert response options were examined with each participant. All participants agreed that the response set (Strongly agree-Strongly disagree) was appropriate for the item stems. Finding an optimal midpoint which portrayed the point between 'Agree' and 'Disagree' proved problematic. There is no conclusive guidance in the literature as to the optimal midpoint label using these response options and some researchers force respondent opinion by excluding a midpoint category (Friedman and Amoo, 1999). Although it can seem desirable to obtain only positive or negative responses, forced choice can introduce bias by indicating a participant has a stronger opinion than they actually do (Friedman and Amoo, 1999). Despite the lack of an optimal solution, Likert scales do have a mid-point and it was therefore felt that including a middle category was desirable. ‘Neither agree nor disagree’ seemed to generate the least controversy among participants. This may be due to it being a common category used in market research and participants were therefore more comfortable with its use.

Stage Two also sought to identify potential technical issues which may be experienced when launching the questionnaire online. Interviewing respondents in their own home or place of work ensured the questionnaire was completed in ‘real life’ settings and tested its use on multiple computers and web browsers. As with paper questionnaires, it was important that respondents receive the questionnaire in a uniform or standardised manner (Baker et al., 2004). Format can be difficult to control in web questionnaires. Computer monitor sizes, for example, can cause the layout to appear different to each respondent. Therefore, the aim was to make the questionnaire appear in the most uniform way possible regardless of the equipment used. BOS software enabled questions to be placed in a grid (matrix) format and provided some uniformity to the design. It was not possible however to fix the width of the grid formatted questions so that questions would appear identically to each participant regardless of the width of the computer monitor. This issue was addressed in subsequent development stages.

When completing the questionnaire, participants did appear to have occasional difficulty navigating from the questionnaire to the external health-related website they were asked to browse and then back to the questionnaire to finish the remaining questions. It was difficult to conclude whether the interviewer’s presence influenced how the participant acted in this instance. It is possible that participants were mindful that the interviewer was present and, however discreetly, was observing their actions. This may have prompted them to ask how they should proceed after browsing the specified website instead of acting how they might have done if alone. In a number of instances when the participant seemed hesitant about how to continue, they were encouraged to complete the questionnaire as they would if they were on their own. With some thought all participants seemed to be able to continue correctly with survey completion. Nonetheless, this was a source of concern and instructions were amended to provide clarity to the participant. It is however noteworthy that, in a setting for which the questionnaire is designed to be used, it is likely that the questionnaire would be integrated fully with the website being assessed making a more participant friendly completion process. For example, if the questionnaire is integrated into a platform which hosts a questionnaire and a prototype website under evaluation, the developers are likely to have more control over the sequence of events (as done in Chapter 10).

Overall Stage Two demonstrated the retained items and their electronic format were suitable for entry into formal psychometric testing (Stage Three) to reduce the number of items.

CHAPTER 6: The pilot eHealth Impact Questionnaire: Item reduction and scale generation

6.1 Introduction

In order to measure a psychological construct, a measurement tool must be easy to administer and have properties which indicate it is both valid and reliable. If the psychometric properties of a measure are not supported, all subsequent analyses may be invalid (Nunnally and Bernstein, 1994). Having constructed a pilot questionnaire relevant to the impact of using health-related websites, it was important to assess the relationships among the items using an adequate sample size. Analyses were primarily led through the use of Classical Test Theory (CTT) which largely requires a scale to contain multiple unidmensional items which are substantially correlated to each other (DeVellis, 2006). In line with CTT, factor analysis was used to reduce the number of items to a more manageable size by removing redundant items, to determine the structure among the items and to identify underlying constructs present within the questionnaire (Norman and Streiner, 2000, Hinkin, 1998).

In addition to refining the questionnaire and identifying sub-scales, it was necessary to establish the consistency with which participants were responding to items within a given scale. The reliability (or internal consistency) of the sub-scales identified were therefore examined using the Cronbach’s alpha statistic which has a recommended value of 0.7 or above (DeVaus, 2002). This value indicates the measure is largely free from random error and responses to items are highly correlated. The item-to-total correlation value was also calculated. This value indicates how strongly an item in a scale correlates to the scale as a whole after omitting that item from the total (Portney and Watkins, 2009). Items with coefficient values of below 0.3 were thought to be tapping into a different variable and were therefore removed to improve the scale's reliability (Kline, 2000). This chapter gives an account of the methods used to refine the two part pilot questionnaire and the process taken to construct theoretically and psychometrically sound sub-scales.

6.1.1 Objectives

Stage Three aimed to administer the pilot questionnaire to an appropriate sample to enable item reduction and psychometric refinement. This was with a view to meeting the following objectives:

- To construct a scale (or sub-scales) relating to general views of using the internet in relation to health (eHIQ-Part 1).
- To construct a scale (or sub-scales) relating to the impact of using a specific health-related website (eHIQ-Part 2).

6.2 Research design

The two part questionnaire was administered online to a range of health groups. To ensure items were suitable for use across a range of patient groups, the sampling framework aimed to reflect a breadth of health conditions. Data were subjected to psychometric analysis to reduce and refine the number of items.

6.3 Methodology

6.3.1 Participants

Men and women who had access to the internet, were aged 18 years or more, lived in the UK and were from one of the following population groups were included:

- People who have asthma
- People who would like to stop smoking
- People who would like to reduce their alcohol intake
- Carers of people with multiple sclerosis (MS)
- Carers of people with motor neurone disease (MND)

These population groups were included to represent the three broad health groups present in all stages of this research (People with long term conditions, people looking for health promotion or screening information and, carers of people with a health condition). Specific conditions within these broad groups were selected as it was expected that they would represent a range of the anticipated themes which it was hypothesised would be transferrable to other similar conditions.

6.3.2 Procedure

6.3.2.1 Recruitment

Open recruitment took place through advertised research invitations on health-related websites such as health blogs, online discussion forums, social networking sites (Facebook and Twitter), news pages on health websites, research volunteer pages, local news online advertisements, research and policy volunteer email list (See Appendix 6.1 for recruitment list and research recruitment advertisements). Recruitment methods were approved by the University of Oxford's Medical Sciences Division Research Ethics Committee (Reference: MSD/IDREC/C1/2011/77) with consent implicit upon completion of the online questionnaire.

6.3.2.2 Sample size

Discussions with several stakeholders suggested that it might prove difficult to recruit participants online. Representatives from the MS Society, for example, thought carers would be hard to reach as they appeared to have less online presence in support forums when compared to people with MS. The aim was to therefore to recruit as many participants as possible within the study timeframe. Estimates suggest that meaningful psychometric tests require at least three times as many respondents as items (Kline, 2000). The longest section of the eHIQ (eHIQ-Part 2) contained 39 items. Therefore 120 participants were required to begin analysis.

6.3.3 Materials

Participant information and instructions (see Appendix 6.2) were kept simple and short as it was expected the questionnaire would be completed by participants with a wide range of abilities and educational levels (Robson, 2002). A separate web-based survey was formatted using BOS software for each population group according to the layout established in Stage Two (BOS). Each survey contained the eHIQ items, demographic items, health literacy items, internet use items and relevant condition specific items (See Table 6-1). All items in addition those in the eHIQ were not considered to be part of the eHIQ, but used to perform descriptive analysis of the sample. To screen for health literacy, the item: How confident are you filling out medical forms yourself? (Wallace et al., 2006). The response category 'somewhat' was used as a cut-off point to identify people with 'limited or marginal health literacy skills'.

Table 6-1 Stage Three survey content

Population group	**Item content**
All participants	
eHIQ-Part 1	General attitudes towards using the internet for health information (23 items)
eHIQ-Part 2	The impact of using a specific health-related website (39 items)
Demographics	Age, Sex, Education, Ethnicity, Employment status
Health literacy	Help reading hospital materials, Confidence filling out medical forms, Problems learning about medical condition because of difficulty understanding written information (Chew et al., 2004)
Internet use	Frequency items adapted adapted from the OII Survey (Blank, 2010)
Population specific	
Asthma	Asthma diagnosis, Length of time since diagnosis, Severity
Smoking cessation	Age when started smoking, Current feelings on smoking
Alcohol reduction	Frequency of drinking alcohol, Amount of alcohol on a typical day when drinking, Frequency of drinking six or more drinks on one occasion (Babor et al., 2001)
Carers (MS and MND)	Length of time being a carer

Upon recruitment, participants were asked to click on an electronic link to the online questionnaire and complete the eHIQ-Part 1. Participants were then asked to spend 10-15 minutes browsing a relevant condition-specific health website, for example a website hosted by: Asthma UK, the MND Association or NHS Choices. Participants were then asked to return to the questionnaire to answer all remaining items. The length of time estimated to complete the questionnaire was 20-25 minutes inclusive of the time allocated to browsing the specified health website.

6.4 Analysis

Raw coded data and an accompanying code book for each survey were downloaded from BOS software to EXCEL. All data were aggregated and imported into a master data file in SPSS (IBM, 2011). All respondents who completed both the eHIQ-Part 1 and the HIQ-Part 2 were included in the subsequent analysis carried out in SPSS. Questionnaire items within the eHIQ-Part 1 and eHIQ-Part 2 had Likert type response options from ‘Strongly agree –Strongly disagree’ which were coded 1-5 where a score of 1 indicated the most negative impact using a website and 5 indicated the most positive impact of using a website. Descriptive analyses were carried out on demographic, condition specific and internet use data. Item reduction and psychometric analysis took place on the eHIQ items.

6.4.1 Item reduction, constructing scales and internal reliability

To ensure the questionnaire was efficient and to minimise responder burden, only relevant items were retained. Items were subjected to preliminary data checks to confirm their suitability for inclusion in further analysis. Floor and ceiling effects may indicate limited scaling properties (Nunnally and Bernstein, 1994, Jenkinson, 1997) and extreme endorsement frequencies are not considered acceptable desirable in the development of psychometric measures as they provide little information (i.e. most people are responding in the same way making the questionnaire less efficient at detecting differences among respondents) (Streiner and Norman, 2008). Therefore, items which had a large number of respondents selecting one of the extreme response options were removed. Items which had large amounts of missing data (> 10% non-response) were also considered inappropriate and removed.

Further item reduction decision steps included using a correlation matrix to identify items demonstrating poor correlations (<0.2) with a large number of items. Poor correlations with a large number of items can indicate a particular item is not tapping a similar construct to items in the scale (i.e. they do not share variance) (Nunnally and Bernstein, 1994, Hinkin, 1998). Reliability analysis identified items with low item-to-total correlations (<0.3) and items which lowered the Cronbach alpha value (Nunnally and Bernstein, 1994). Items displaying a high number of poor correlations with other items or items which lowered the Cronbach's alpha value were iteratively removed.

6.4.2 Factor analysis

Factor analysis was carried out to identify the number of sub-scales (or factors) within the eHIQ-Part 1 and the eHIQ-Part 2 (DeVellis, 2006). Although the secondary qualitative analysis (Chapter 4) had identified five themes of relevance, no a priori assumptions were made regarding how the items would group together when subjected to quantitative analyses. It was therefore considered appropriate to subject remaining items to exploratory factor analysis to determine how items quantitatively grouped together (Norman and Streiner, 2000, Matsunaga, 2010). This process sought to identify sets of items which showed the most evidence of similarity (or shared variance) until there was virtually no meaningful association left between the remaining items (DeVellis, 2006).

The suitability of the dataset for factor analysis was assessed. Item correlation matrices were tested for a significant Bartlett's Test of Sphericity ($p<0.05$) (Bartlett, 1954). Bartlett's Test of Sphericity performs chi-square tests on the item correlation matrix to establish if there is a relationship among the items (Pett et al., 2003). The sample was tested for appropriateness of factor analysis using the Kaiser-Meyer-Olkin (KMO) value (recommended value of 0.6) (Kaiser and Rice, 1974). The KMO statistic is a measure of sampling adequacy which compares the magnitude of correlation coefficients to partial correlation coefficients between items (Pett et al., 2003). Partial correlations account for the correlation between items due to a common underlying factor and cancel out the effects of other items. If items share common factors, the partial correlations between two items should be small when the effects of other items have been removed (Norman and Streiner, 2000, Pett et al., 2003). Further inspection of the dataset to determine its suitability for factor analysis involved examining the anti-image correlation matrix (AIC) output from SPSS. The AIC matrix compares the negative partial correlation between items (i.e. the correlation between variables after the effect of other variables has been removed) and the measure of sampling adequacy (MSA) for each individual item. The MSA indicates how strongly an item is related to other items in the matrix. It is important to ensure the values of negative partial correlations are small and the MSA values are large (Pett et al., 2003). Communalities (squared multiple correlations) were also examined to ensure the items were related to each other (value over 0.6) (Norman and Streiner, 2000).

Items were subjected to factor analysis within SPSS and factors were assessed using a number of steps. Factors with Eigenvalues above one were initially extracted. A factor's Eigenvalue indicates how much of the total variance of the entire measure that particular factor can explain. Some regard this set cut off point as somewhat arbitrary as, for example, factors with an Eigenvalue of 0.99 should not always be automatically disregarded (Norman and Streiner, 2000). During this developmental stage, Cattell's Scree test was also used to determine what constituted a meaningful factor (Cattell, 1966, Norman and Streiner, 2000). A Scree plot can be computed within SPSS to show a graphical representation of the Eigenvalues to identify a sharp break in the descending curve (before the scree begins). Where it was difficult to make conclusive judgments from the scree plot (i.e. difficult to ascertain where the scree began), the items loadings were assessed from a theoretical perspective and intuitive judgments were made about how many factors to retain (Pett et al., 2003).

Factors were rotated to find clusters of items which shared the same strong characteristics and aid interpretation by spreading the variance more evenly across factors (Nunnally and Bernstein, 1994, DeVellis, 2006). An oblique solution was sought using Direct Oblimin rotation as it was felt that factors should not be restricted to independent factors (as with orthogonal rotations). Oblique rotations do not restrict axes to right angles and can therefore be preferable as they allow some correlation between the factors (Norman and Streiner, 2000, Matsunaga, 2010). Furthermore, oblique rotations do not rule out the possibility of an orthogonal solution where it is the optimal solution.

Two matrices of interest are produced following an oblique rotation: the Structure matrix and the Pattern matrix. The Structure matrix loadings represent the correlations of the items with the factors whilst the Pattern matrix loadings represent the effect of a given factor on a given item when other factors are controlled. Pattern loadings are similar to partial standardized regression coefficients in a regression (Pett et al., 2003). Whilst both matrices were taken into account when interpreting the factor rotation, the Structure matrix offered primary guidance in interpretation and then compared to the Pattern loadings for consistency. Several authors (Pett et al., 2003, Nunnally and Bernstein, 1994, Kline, 1994) support the use of the Structure matrix as its values are more stable compared to the pattern loadings which are affected by how correlated the factors are to each other.

Internal consistency was tested for each unidimensional scale identified during factor analysis using the Cronbach's alpha statistic (>0.7) (Nunnally and Bernstein, 1994). Checks for evidence of further scales present were made on all items removed to ensure that important items were not discarded or important issues were not covered.

6.5 Results

This section begins with a description of the population sampled. Stage Three aimed to reduce the eHIQ-Part 1 and the eHIQ-Part 2 to a more manageable number of items and identify the underlying sub-scales. The results of preliminary item reduction and subsequent factor analysis are reported in detail for each questionnaire part.

6.5.1 Characteristics

The two part questionnaire was completed online by 167 participants. The mean age of the sample was 38.8 years (SD 14.04) and just over half the participants (n=92, 55.1%) had a long term condition (See Table 6-2). Forty-one (24.6%) male respondents took part in the study with women outnumbering men in each health group. Of those who responded, 152 (94.4%) said they had previous experience of visiting a health website and 97.6% (n=163) had completed secondary level education. A small number of participants reported poor health literacy with nine participants (5.5%) reporting limited or marginal health literacy skills as defined by the ‘somewhat’ response (see Table 6-3).

Of the 17 participants in the smoking cessation sample, 15 participants had given up smoking recently or had some desire to stop smoking (see Table 6-4). Participants who indicated that they had given up smoking recently were included as the website viewed was also relevant to people who had recently stopped smoking and offered tips on how to refrain from taking up smoking again. Eleven of the 25 participants in the alcohol reduction sample drank alcohol at least twice a week and 76 of the 92 people in the asthma sample considered their asthma to be moderate to severe. Those caring for people with MND had been caring for an average of 4.1 years while those caring for someone with MS averaged 15.0 years caring experience. It is worth noting that the average length of time for those caring for people with MND was slightly longer than expected given a median survival rate of approximately 3-4 years from onset of symptoms (Baumer et al., 2014). This may be due to the small sample size, but may also indicate that those caring for longer are more likely to take part in research. Furthermore there is an increasing belief that time with a diagnosis of MND is longer than once thought, due in part to improvements in diagnosis (Baumer et al., 2014).

Table 6-2 Stage Three age by health condition

Condition	n	Mean age (SD)
Smoking cessation	17	37.53 (13.33)
Alcohol reduction	25	32.52 (9.67)
Motor Neurone Disease (carer)	11	51.55 (12.35)
Multiple Sclerosis (carer)	22	48.45 (15.49)
Asthma	92	36.91 (13.22)
Total	167	38.80 (14.04)

Table 6-3 Stage Three participant characteristics

Characteristic	Male	Female	Total
Condition			
Smoking cessation	5	12	17
Alcohol reduction	11	14	25
Motor Neurone Disease (carer)	3	8	11
Multiple Sclerosis (carer)	7	15	22
Asthma	15	76	91
Total	41	125	166
Missing			1
Frequency of visits to health websites			
At least once a week	16	50	66
Monthly	10	34	44
Less than monthly	10	32	42
Never	3	3	6
Don't know	2	1	3
Total	41	120	161
Missing			6
Employment status			
At work	21	52	73
Unemployed	2	3	5
Student	5	19	24
Not working (ill health)	2	23	25
Home duties	1	8	9
Retired (ill health or old age)	8	11	19
Other	2	8	10
Total	41	124	165
Missing			2
Health literacy			
Limited or marginal health literacy skills	3	6	9

Table 6-4 Stage Three condition specific participant characteristics

Condition group	Frequency
Smoking cessation	
Smoking status	
No desire to stop smoking	2
Intention to stop smoking	9
Recently stopped smoking	4
Other	2
Total	17
Alcohol reduction	
Drink containing alcohol	
Monthly or less	5
2-4 times a month	9
2-3 times a week	7
4 or more times a week	4
Total	25
Asthma	
Severity	
Mild	16
Moderate	42
Severe	34
Total	92
Carers (MND)	
Mean length of time caring Years(Standard deviation)	4.1 (SD 4.1)
Carers (MS)	
Mean length of time caring Years(Standard deviation)	15.0 (SD 9.1)

6.5.2 Stage Three pilot questionnaire psychometric properties

6.5.2.1 Floor and ceiling effects and missing data

Frequencies of selected response categories for all eHIQ items were calculated. Checks for floor and ceiling effects confirmed no item had more than 40% of respondents selecting one of the extreme response options. Checks also confirmed no item had a large amount of missing data with all ≤ 4.2% non-response. Therefore, no items were removed due to floor or ceiling effects or missing responses. All subsequent analyses for the eHIQ-Part 1 and eHIQ-Part 2 were carried out separately.

6.5.2.2 Preliminary item reduction (eHIQ-Part1)

Seven items were identified and removed due to having poor correlations (<0.2) with a large number of items within the correlation matrix and due to lowering the Cronbach's alpha value. Items were iteratively removed by sequentially identifying and removing the worst offending items. After items were removed, a new correlation matrix was produced and reliability analysis rerun at each iteration (see Table 6-5 for each iteration and items removed).

Table 6-5 Stage Three Preliminary item reduction (eHIQ-Part 1)

Iteration	Total number of items	Cronbach's alpha value	Potential items for deletion	Number of poor correlations with other items (<0.2)	Item reducing Cronbach's alpha value	Item removed
1	23	0.83	7) People become isolated from everyday life when they use health websites extensively.	14	✓	✗
			9) I prefer to avoid websites that include images of health problems (For example, a rash on a person's skin).	15	✓	✗
			13) When I look at someone's health experience on a website I like to see what the person looks like.	17	✓	✓
			14) I worry about all the horror stories on health websites.	15	✓	✓
			19) I worry about people making negative personal comments on health websites.	18	✓	✓
2	20	0.84	1) The internet is a good way of finding other people who are experiencing similar health problems.	10	✗	✗
			6) It can be easier to share health concerns with a stranger on the internet than with those around you.	11	✗	✗
			7) People become isolated from everyday life when they use health websites extensively.	11	✓	✓
			9) I prefer to avoid websites that include images of health problems (For example, a rash on a person's skin).	12	✓	✓
3	18	0.85	1) The internet is a good way of finding other people who are experiencing similar health problems.	9	✓	✗
			6) It can be easier to share health concerns with a stranger on the internet than with those around you.	10	✓	✓
4	16	0.85	2) It is difficult to find reliable health information on the internet.	8	✓	✓

6.5.2.3 Factor analysis (eHIQ-Part 1)

The 16 remaining items were assessed for their suitability to factor analysis. The KMO value was 0.82, exceeding the recommended value of 0.6 and the Bartlett's Test of Sphericity reached statistical significance ($p<0.01$), supporting the use of factor analysis.

Five factors with Eigenvalues above one were initially extracted. These factors explained 63.85% of the variance. However, Catell's Scree test suggested that only two factors should be extracted. Conceptually, factors 3-5 were not believed to constitute meaningful factors and their removal was further supported by each factor having a Cronbach's alpha of below 0.7. Seven items which loaded on factors 3-5 were therefore removed and are listed in Table 6-6, (Iteration 1).

Upon the removal of the seven items above, factor analysis was performed again and a further item was removed due to poor communality (Table 6-6, Iteration 2). Item 22 had a low communality (0.35) and loaded weakly on all factors. This may have been due to the slightly ambiguous wording of the item. The remaining items were entered into a final factor analysis resulting in two factors explaining 61.06% of the total variance.

Table 6-6 Stage Three Items removed during factor analysis (eHIQ-Part 1)

Factor analysis iteration	Reason for removal	Item removed
1	Seven items which loaded on factors 3-5 (not believed to constitute meaningful factors)	Q11) I can work out what health information to trust on the internet.
		Q12) The internet is a good resource to help me understand what a doctor tells me.
		Q15) I can usually find relevant health information on the internet.
		Q17) I feel my health experiences could be of value to other people on the internet.
		Q18) I think I can tell if a person's story on a health website sounds accurate.
		Q21) The internet causes people to worry unnecessarily about their health.
		Q23) The internet is a good resource which I can use to check the doctor is giving me appropriate advice.
2	Poor communality	Q22) Health websites can help the public to know what it is like to live with a health problem.

6.5.2.4 Factor rotation (eHIQ-Part 1)

The two factors consisting of eight items were subjected to an oblique, Direct Oblimin rotation as it was hypothesised that factors should correlate. The Structure matrix (see Table 6-7) and the Pattern matrix (see Table 6-8) output were used to interpret the factor structure. Despite the structure matrix offering primary guidance, it was noted that the pattern matrix loadings allocated the same items to each factor. Internal consistencies for both factors were good (Cronbach’s alpha, Factor 1= 0.82 and Factor 2= 0.76).

Table 6-7 Stage Three Structure Matrix (eHIQ-Part 1)

Item	Factor	
	1	2
3) It can be helpful to see other people's health experiences on the internet.	0.81	0.18
1) The internet is a good way of finding other people who are experiencing similar health problems.	0.80	0.26
16) The internet is a good way of finding other people who are facing health related decisions I may also face.	0.74	0.44
4) The internet is useful if you don't want to tell people in your everyday life how you feel.	0.70	0.21
5) It can be reassuring to know that I can access health websites any time of the day or night.	0.67	0.32
20) Looking at health websites reassures me that I am not alone with my health concerns.	0.61	0.56
8) The internet can be useful to help people decide if their symptoms are important enough to go and see a doctor.	0.32	0.89
10) I would use the internet if I needed help to make a decision about my health (For example, whether I should seek medical advice or seek alternative therapies).	0.32	0.88

Extraction Method: Principal Component Analysis. Rotation Method: Oblimin with Kaiser Normalization.

Table 6-8 Stage Three Pattern Matrix (eHIQ-Part 1)

Item	Factor	
	1	2
3) It can be helpful to see other people's health experiences on the internet.	0.87	-0.15
1) The internet is a good way of finding other people who are experiencing similar health problems.	0.81	-0.04
4) The internet is useful if you don't want to tell people in your everyday life how you feel.	0.72	-0.05
16) The internet is a good way of finding other people who are facing health related decisions I may also face.	0.67	0.19
5) It can be reassuring to know that I can access health websites any time of the day or night.	0.64	0.08
20) Looking at health websites reassures me that I am not alone with my health concerns.	0.47	0.38
8) The internet can be useful to help people decide if their symptoms are important enough to go and see a doctor.	-0.02	0.90
10) I would use the internet if I needed help to make a decision about my health (For example, whether I should seek medical advice or seek alternative therapies).	-0.01	0.88

Extraction Method: Principal Component Analysis. Rotation Method: Oblimin with Kaiser Normalization.

6.5.2.5 Re-analysing removed items

As Stage Three was a development stage of this research, further consideration was given to the items removed during analysis. In the interests of ensuring items had not been unduly disregarded, they were examined to see if they constituted a standalone scale. Six of the eight items achieved a Cronbach's alpha value of 0.64, slightly below the recommended value of 0.7. The six items were retained however for the next stage of analysis after considering their theoretical value. The remaining two items (Q17: I feel my health experiences could be of value to other people on the internet and Q18: I think I can tell if a person's story on a health website sounds accurate) were removed due to poor item-total values.

6.5.3 Items retained for Stage Four (eHIQ-Part 1)

In summary, fourteen items asking about participants general attitudes towards health-related websites were identified for inclusion in eHIQ-Part 1. The three sub-scales grouped around: 1) Attitudes towards sharing health experiences online, 2) Use of the internet for decision-making, and 3) Attitudes towards online health information in general. All items are shown in Table 6-9.

Table 6-9 Stage Three items grouped by sub-scale (eHIQ-Part 1)

Sub-scale	Items	Item to total correlation	Cronbach's alpha
1) Items relating to attitudes towards sharing health experiences online *(6 items)*	1) The internet is a good way of finding other people who are experiencing similar health problems	0.63	0.82
	2) It can be helpful to see other people's health experiences on the internet.	0.66	
	3) The internet is useful if you don't want to tell people in your everyday life how you feel.	0.54	
	4) It can be reassuring to know that I can access health websites any time of the day or night.	0.53	
	10) The internet is a good way of finding other people who are facing health related decisions I may also face.	0.62	
	11) Looking at health websites reassures me that I am not alone with my health concerns.	0.51	
2) Items relating to decision-making *(2 items)*	5) The internet can be useful to help people decide if their symptoms are important enough to go and see a doctor.	0.62	0.76
	6) I would use the internet if I needed help to make a decision about my health. (For example, whether I should seek medical advice or seek alternative therapies)	0.62	
3) Items relating to attitudes towards online health information *(Standalone 6 item scale)*	7) I know what health information to trust on the internet.	0.35	0.64
	8) The internet is a reliable resource to help me understand what a doctor tells me.	0.32	
	9) I know how to find relevant health information on the internet.	0.45	
	12) The internet causes people to worry about their health.	0.35	
	13) The internet can help the public to know what it is like to live with a health problem.	0.35	
	14) The internet is a resource which I can use to check the doctor is giving me appropriate advice.	0.44	

6.5.3.1 Preliminary item reduction (eHIQ-Part 2)

Five items were identified and removed due to having poor correlations (<0.2) with a large number of items in the correlation matrix and due to reducing the overall Cronbach's alpha value. Two of the five candidate items for removal were also thought to be similar to remaining items (Q19 similar to Q24, and Q4 similar to Q17 and Q8) which supported the decision for their removal. Items were iteratively removed (See Table 6-10) by sequentially identifying and removing the worst offending items. After each iteration, a new correlation matrix was produced and reliability analysis rerun at each iteration. Five items were removed during three iterations. No further items were removed following the fourth iteration there was a relatively low number of poor correlations with other items.

Table 6-10 Stage Three Preliminary item reduction table (eHIQ-Part 2)

Iteration	Total number of items	Cronbach's alpha value	Potential items for deletion	Number of poor correlations with other items (<0.2)	Item reducing Cronbach's alpha value	Items removed
1	39	0.95	4) This website seems to be aimed at a different age group to my needs.	19	✓	✗
			5) This website raises questions I might ask a doctor or nurse.	16	✓	✗
			16) This website makes me worry about things that are out of my control.	24	✓	✓
			19) I found the information on this website distressing.	16	✗	✗
			31) This website makes me worry that I may not receive the health care I need.	20	✓	✓
2	37	0.95	4) This website seems to be aimed at a different age group to my needs.	17	✓	✓
			5) This website raises questions I might ask a doctor or nurse.	10	✗	✗
			19) I found the information on this website distressing.	15	✗	✓
			37) I would not want someone I care for to see this website.	7	✗	✗
3	35	0.95	3) The information on this website left me feeling confused.	6	✗	✗
			5) This website raises questions I might ask a doctor or nurse.	8	✓	✓
			11) I feel more inclined to look after myself after visiting this website.	6	✗	✗
			14) This website provides more useful information than my doctor is likely to give me.	6	✗	✗
			37) I would not want someone I care for to see this website.	7	✓	✗
4	34	0.95	3) The information on this website left me feeling confused.	5	✗	✗
			11) I feel more inclined to look after myself after visiting this website.	5	✗	✗
			14) This website provides more useful information than my doctor is likely to give me.	*5*	✗	✗

6.5.3.2 Factor analysis (eHIQ-Part 2)

The 34 remaining items were assessed for their suitability for factor analysis. The KMO value was 0.92, exceeding the recommended value of 0.6 and the Bartlett's Test of Sphericity reached statistical significance (p<0.01), supporting the use of factor analysis. Six factors with Eigenvalues above one were initially extracted. These factors explained 66.26% of the variance. Catell's Scree test was unclear, therefore the six factors were extracted for further investigation.

6.5.3.3 Factor rotation (eHIQ-Part 2)

The remaining six factors were subjected to a Direct Oblimin rotation as it was thought that that they were likely to correlate. Following rotation, the Structure matrix (see Table 6-11) and Pattern matrix (see Table 6-12) were used to interpret the output. As with the loadings for eHIQ-Part 1, the Structure matrix offered the primary source of guidance for the factor structure and the Pattern matrix showed little difference in factor loadings. Item 15 loaded highly on Factors 1 and 5 in the Structure and Pattern matrices. It was felt however that this item fitted conceptually with items in Factor 1 more so than with items in Factor 5. All factors had a Cronbach's alpha value of 0.76 or above.

Table 6-11 Stage Three Structure Matrix (eHIQ-Part 2)

Item	Factor					
	1	2	3	4	5	6
11) I feel more inclined to look after myself after visiting this website.	0.81	0.03	-0.31	0.19	0.30	-0.24
35) This websites gives me the confidence to discuss my health with the people in my life (For example, my family or people at work).	0.81	0.13	-0.39	0.28	0.36	-0.54
12) I have learnt something new from this website.	0.80	0.18	-0.25	0.20	0.36	-0.44
34) This website makes me want to play a more active role in my healthcare.	0.79	0.16	-0.37	0.35	0.38	-0.49
33) This website helped me to develop a better understanding of my personal health.	0.76	0.25	-0.48	0.24	0.44	-0.55
15) This website prepares me for what might happen to my health.	0.72	0.23	-0.39	0.35	0.71	-0.49
32) This website gives me the confidence to explain my health concerns to others.	0.70	0.20	-0.38	0.22	0.25	-0.59
19) I wish I had seen this website at an earlier stage.	0.65	0.04	-0.24	0.19	0.46	-0.49
29) This website gives me confidence that I am able to manage my health.	0.65	0.33	-0.45	0.34	0.49	-0.39
37) I found the images on this website distressing.	0.08	0.84	-0.33	0.17	0.11	-0.23
27) I would not want someone I care for to see this website.	0.19	0.83	-0.34	0.30	0.32	-0.18
24) On the whole, I found this website depressing.	0.22	0.78	-0.33	0.34	0.02	-0.32
3) The information on this website left me feeling confused.	0.01	0.72	-0.39	0.47	0.30	-0.15
17) The information on this website does not apply to me.	0.14	0.59	-0.08	0.28	0.43	-0.45
39) This website is easy to use.	0.33	0.30	-0.77	0.30	0.16	-0.17
38) This website is attractive.	0.32	0.20	-0.76	0.30	0.26	-0.17

36) Photographs and other images were used appropriately on the website.	0.35	0.33	-0.75	-0.04	0.23	-0.33
13) I can easily understand the information on this website.	0.44	0.41	-0.74	0.43	0.46	-0.39
9) The language on this website made it easy to understand.	0.31	0.39	-0.74	0.36	0.31	-0.39
8) This website gives a good range of information.	0.32	0.33	-0.68	0.53	0.35	-0.45
10) This website left me with a lot of unanswered questions.	0.26	0.53	-0.34	0.76	0.37	-0.31
2) This website has a positive outlook.	0.44	0.45	-0.48	0.66	0.22	-0.21
6) This website includes useful tips on how to make life better.	0.45	0.12	-0.45	0.65	0.24	-0.33
7) This website seems patronising.	0.15	0.47	-0.14	0.64	0.15	-0.47
1) This website encourages me to take steps that could be beneficial to my health.	0.57	0.17	-0.52	0.59	0.40	-0.18
14) This website provides more useful information than my doctor is likely to give me.	0.53	-0.03	-0.18	0.18	0.74	-0.36
21) I trust the information on this website.	0.27	0.33	-0.60	0.31	0.67	-0.14
22) I would consult this website if I had to make a decision about my health.	0.55	0.29	-0.40	0.10	0.67	-0.45
28) I value the advice given on this website.	0.61	0.43	-0.49	0.34	0.64	-0.44
30) I feel I have a lot in common with other people using this website.	0.55	0.28	-0.32	0.26	0.31	-0.86
23) I feel I have a sense of solidarity with other people using this website.	0.59	0.20	-0.39	0.27	0.43	-0.83
25) I can identify with other people using this website.	0.49	0.27	-0.41	0.27	0.32	-0.81
18) The contributors to this website understand what is important to me.	0.60	0.30	-0.49	0.53	0.47	-0.68
26) On the whole, I find this website reassuring.	0.43	0.37	-0.35	0.38	0.57	-0.61

Extraction Method: Principal Component Analysis. Rotation Method: Oblimin with Kaiser Normalization.

Table 6-12 Stage Three Pattern Matrix (eHIQ-Part 2)

Item	Factor					
	1	2	3	4	5	6
11) I feel more inclined to look after myself after visiting this website.	0.86	-0.03	-0.02	0.02	0.02	0.14
12) I have learnt something new from this website.	0.78	0.12	0.11	-0.02	0.05	-0.08
34) This website makes me want to play a more active role in my healthcare.	0.67	0.00	-0.03	0.14	0.05	-0.14
35) This websites gives me the confidence to discuss my health with the people in my life (For example, my family or people at work).	0.67	-0.04	-0.08	0.06	0.02	-0.22
33) This website helped me to develop a better understanding of my personal health.	0.56	0.08	-0.17	-0.04	0.11	-0.22
32) This website gives me the confidence to explain my health concerns to others.	0.55	0.05	-0.11	-0.01	-0.10	-0.34
19) I wish I had seen this website at an earlier stage.	0.47	-0.11	0.03	0.01	0.24	-0.25
29) This website gives me confidence that I am able to manage my health.	0.47	0.17	-0.13	0.09	0.22	-0.02
37) I found the images on this website distressing.	-0.01	0.84	-0.13	-0.11	-0.08	-0.03
27)I would not want someone I care for to see this website	0.10	0.81	-0.05	0.02	0.14	0.13
24) On the whole, I found this website depressing.	0.17	0.75	-0.06	0.10	-0.26	-0.09
3) The information on this website left me feeling confused.	-0.20	0.59	-0.17	0.27	0.17	0.07
17) The information on this website does not apply to me.	-0.09	0.49	0.21	0.08	0.31	-0.30
36) Photographs and other images were used appropriately on the website.	0.08	0.16	-0.74	-0.36	-0.01	-0.16
38) This website is attractive.	0.03	-0.05	-0.73	0.09	0.06	0.04
39) This website is easy to use.	0.09	0.08	-0.72	0.08	-0.09	0.06

9) The language on this website made it easy to understand.	-0.07	0.11	-0.64	0.09	0.06	-0.21
13) I can easily understand the information on this website.	0.07	0.13	-0.56	0.14	0.19	-0.11
8) This website gives a good range of information.	-0.12	-0.02	-0.55	0.32	0.10	-0.28
10) This website left me with a lot of unanswered questions.	0.00	0.30	-0.01	0.62	0.16	-0.04
6) This website includes useful tips on how to make life better.	0.21	-0.16	-0.24	0.56	-0.03	-0.10
7) This website seems patronising.	-0.09	0.28	0.15	0.55	-0.06	-0.37
2) This website has a positive outlook.	0.31	0.26	-0.19	0.49	-0.07	0.12
1) This website encourages me to take steps that could be beneficial to my health.	0.39	-0.06	-0.26	0.45	0.15	0.18
14) This website provides more useful information than my doctor is likely to give me.	0.28	-0.19	0.08	0.02	0.66	-0.10
21) I trust the information on this website.	-0.08	0.10	-0.45	0.07	0.59	0.15
22) I would consult this website if I had to make a decision about my health.	0.27	0.14	-0.14	-0.20	0.49	-0.16
15) This website prepares me for what might happen to my health.	0.45	0.02	-0.03	0.09	0.47	-0.12
28) I value the advice given on this website.	0.34	0.24	-0.15	0.04	0.40	-0.06
30) I feel I have a lot in common with other people using this website.	0.21	0.04	-0.05	0.02	-0.02	-0.75
25) I can identify with other people using this website.	0.10	0.01	-0.19	0.03	0.01	-0.71
23) I feel I have a sense of solidarity with other people using this website.	0.20	-0.07	-0.14	0.02	0.12	-0.69
18) The contributors to this website understand what is important to me.	0.22	-0.01	-0.18	0.30	0.15	-0.43
26) On the whole, I find this website reassuring.	0.05	0.13	-0.06	0.14	0.36	-0.40

Extraction Method: Principal Component Analysis. Rotation Method: Oblimin with Kaiser Normalization

6.5.4 Items retained for Stage Four (eHIQ- Part 2)

Thirty-four items asking about using a specific health-related website were identified as suitable to enter the next stage of development (see Table 6-13). The six sub-scales related to: 1) Items relating to managing health, 2) Items relating to apprehension and distress over website content, 3) Items relating to presentation of information, 4) Items relating to the tone of a website, 5) Items relating to trust and value of website content, and finally 6) Items relating to identifying with website content.

Table 6-13 Stage Three Items grouped by sub-scales (eHIQ-Part 2)

Sub-scale	Items	Item to total correlation	Cronbach's alpha
1) Items relating to managing health *(9 items)*	11) I feel more inclined to look after myself after visiting this website.	0.66	0.92
	35) This websites gives me the confidence to discuss my health with the people in my life (For example, my family or people at work).	0.78	
	12) I have learnt something new from this website	0.69	
	34) This website makes me want to play a more active role in my healthcare.	0.78	
	33) This website helped me to develop a better understanding of my personal health.	0.78	
	15) This website prepares me for what might happen to my health.	0.72	
	32) This website gives me the confidence to explain my health concerns to others.	0.68	
	19) I wish I had seen this website at an earlier stage.	0.64	
	29) This website gives me confidence that I am able to manage my health.	0.66	
2) Items relating to apprehension and distress over website content *(5 items)*	37) I found the images on this website distressing.	0.65	0.82
	27) I would not want someone I care for to see this website.	0.71	
	24) On the whole, I found this website depressing.	0.66	
	3) The information on this website left me feeling confused.	0.64	

	17) The information on this website does not apply to me.	0.46	
3) Items relating to presentation of information *(6 items)*	39) This website is easy to use.	0.64	0.85
	38) This website is attractive.	0.60	
	36) Photographs and other images were used appropriately on the website.	0.62	
	13) I can easily understand the information on this website.	0.66	
	9) The language on this website made it easy to understand.	0.66	
	8) This website gives a good range of information.	0.61	
4) Items relating to the tone of a website *(5 items)*	10) This website left me with a lot of unanswered questions.	0.59	0.78
	2) This website has a positive outlook.	0.62	
	6) This website includes useful tips on how to make life better.	0.54	
	7) This website seems patronising.	0.51	
	1) This website encourages me to take steps that could be beneficial to my health.	0.54	
5) Items relating to trust and value of website content *(4 items)*	14) This website provides more useful information than my doctor is likely to give me.	0.51	0.76
	21) I trust the information on this website.	0.52	
	22) I would consult this website if I had to make a decision about my health.	0.61	
	28) I value the advice given on this website.	0.69	
6) Items relating to identifying with website content *(5 items)*	30) I feel I have a lot in common with other people using this website.	0.79	0.89
	23) I feel I have a sense of solidarity with other people using this website.	0.81	
	25) I can identify with other people using this website.	0.79	
	18) The contributors to this website understand what is important to me.	0.72	
	26) On the whole, I find this website reassuring.	0.59	

6.6 Discussion

It is important to establish the psychometric properties of any health-related measurement tool. In order to have confidence that the instrument is measuring what it purports to measure and that it will do so in a consistent basis, it must go through several stages of development. Stage Three of this body of research aimed to reduce the eHIQ to a manageable number of items and examine the underlying constructs. Using psychometric techniques associated with Classical Test Theory, nine items were removed from the eHIQ-Part 1 and five items were removed from the eHIQ-Part 2. The eHIQ-Part 1 asks about the general views of using the internet in relation to health and contains three tentative sub-scales (14 items). The eHIQ-Part 2 asks about the impact of using a specific health-related website and contains six tentative sub-scales (34 items). All sub-scales were conceptually sound in relation to using the internet for health information and were considered tentative as they would undergo further analysis in Stage Four.

Internal reliability was found to be acceptable for all sub-scales. One sub-scale in the eHIQ-Part 1 was constructed as a standalone scale (Cronbach's alpha =0.64). This value is considered acceptable by some standards in psychometric testing (Helmstadter, 1964). Stage Three was considered a developmental stage and the items within the stand alone scale were conceptually relevant, it was decided to retain the items for further examination in the subsequent round of questionnaire analysis.

It was essential that candidate items were tested among a range of population groups and health roles (i.e. patients and carers). The recruitment strategy in this study therefore ensured it targeted people with long term conditions, carers and those considering changes in lifestyle or facing health decisions. Due to the nature of open recruitment on the internet (i.e. advertising an unrestricted access link), it was not possible to estimate response rates or ascertain if a representative sample was achieved. It is possible that the online sampling strategy meant only a subset of the population was targeted (Dillman, 2011). An alternative method of recruitment could have been to obtain a random sample through non-internet means and ask participants to complete the questionnaire (Dillman, 2011). This mode of recruitment however would have also had methodological limitations. For example, unless those who did not have access to the internet were given internet access, recruitment in this manner would still achieve an unrepresentative sample of the general population. The overriding principle supporting recruitment online in this study was connected to the research question and field of research. It was not appropriate to include people who did not have access to the internet in this study as any intervention being assessed with this measure would only be suitable for people who use the internet. Nonetheless, it was thought that alternative methods of recruitment should be explored in the next stage of the study.

In the early stages of administration, it was necessary to add more user friendly features to the online questionnaire. Further navigational instructions and survey completion updates (for example, ‘You have now completed one of three sections of this questionnaire’) were added. Demonstrating survey progress was somewhat limited by the features available on the survey software (i.e. no progress bar option available). During the early stages of administration, two carers also provided feedback via email stating that they were unsure whether they should answer a number of items from the perspective of their own health or the health of the person they cared for. An additional preamble was subsequently included to make it clearer that items related to the participant's own health or experience.

A further problem encountered during Stage Three was the drop out or non-completion rate. There was substantial drop out between completing the eHIQ-Part 1 and the eHIQ-Part 2. It is difficult to ascertain the reasons behind drop out however reasons may typically include difficulties with item completion, difficulties with questionnaire format, technical problems (such as a broken questionnaire link or the effects of technological equipment) (DeVaus, 2002). Despite taking substantial care to ensure detailed instructions were included throughout the questionnaire, a minimum amount of computer proficiency was required to successfully complete the questionnaire (Dillman, 2011). It was largely suspected that a substantial amount of the drop out between the completion of eHIQ-Part 1 and commencing eHIQ-Part 2 was due to asking the participant to click on a link to an external health-related website and then require them to navigate back to the questionnaire page, located on a different tab, to complete the remaining items. Although somewhat out of the researchers' control, these concerns were taken forward for consideration in the subsequent stage of administration. Further efforts to standardise format and experience of completing questionnaire are taken to diminish these potential problems in the next stage of this research.

The next stage of this body of research therefore sought to further improve the electronic format of the online questionnaire and complete further investigations into its psychometric properties.

6.7 Conclusion

In summary, Stage Three was an important step in reducing the number of items within each questionnaire part. It highlighted a number of potential difficulties with the electronic questionnaire in its current format. In addition to carrying out further investigations into the psychometric properties of the remaining items, it was important to address the technical issues that arose in the next stage.

CHAPTER 7: The eHealth Impact Questionnaire: Construct validity, test-retest reliability and translatability

7.1 Introduction

The primary aims of this chapter are to report the final eHIQ sub-scales and outline the research undertaken to establish the psychometric properties of the new questionnaire. This chapter gives an account of the research carried out to support the validity and reliability of the questionnaire across a range of conditions, the potential translatability of the questionnaire into other languages, and explores the merits of various recruitment strategies for the online questionnaire.

As discussed in the previous chapter, good internal reliability (or internal consistency) indicates a scale is free from random error and that responses to items are highly correlated. In addition to confirming the internal consistency of each scale, best practice encourages the assessment of the stability of scores over time (test-retest reliability) (Kline, 2000, Reeve et al., 2013, FDA, 2009). Test-retest reliability stipulates that, providing there has been no underlying change in the construct being measured, an instrument should produce the same (or very similar) scores when completed by a participant on two or more occasions (Jenkinson and McGee, 1997, Hays and Revicki, 2005). To determine test-retest reliability, a questionnaire can be administered on two occasions with a time interval that is sufficiently distant so that the respondent cannot simply replicate their answers from memory and short enough so that the underlying construct has not changed for that respondent (Streiner and Norman, 2008). An instrument can appear to have poor test-retest reliability for a number of reasons. These include: the construct being assessed has changed, contextual issues have changed (for example, the participant feels ill, tired, bored or anxious about an unrelated event), the respondent has been influenced by the process of completing the questionnaire on the first occasion, or the measure is unreliable (Streiner and Norman, 2008, Kline, 2000). The Intra Class Coefficient (ICC) statistic is recommended for calculating test-retest reliability within a given population (Streiner and Norman, 2008). Scale scores for each participant on the first and second completion are correlated and values of greater than 0.7 (or slightly lower at the group level) indicate good test-retest reliability (Kline, 2000).

Testing the reliability of a questionnaire is essential to ascertain the precision of measurement. Establishing precision however cannot meaningfully inform the user if the instrument is in fact not measuring the intended construct. Various forms of validity must therefore be established during a questionnaire's development. Content and face validity were demonstrated in Stages One and Two. Having reduced the number of items within the questionnaire in Stage Three, Stage Four sought to establish construct validity. This can typically be achieved through administering the new questionnaire together with an established reference measure which is expected to have moderate correlations with identified sub-scales. This tests an aspect of construct validity, known as convergent validity. Convergent validity tests an apriori hypothesis that a relationship exists between the new scale(s) and the established 'reference' measure (Hinkin, 1998, FDA, 2009).

Criterion validity can also support construct validity through administering a gold standard measure together with the new measure and subsequently assessing their correlation. Discounting the development of a shorter version of the original instrument however, there are rarely gold standard measures available which measure the construct of interest. As there was no gold standard to assess the impact of using health-related websites, it was not possible or appropriate to evaluate criterion validity at this stage.

During the development of a new measure it can also be useful to assess the translatability of items so that minor amendments can be made to item wording before the final data collection takes place. Translatability Assessments evaluate the extent to which a measure can be meaningfully translated into another language so that it is both conceptually equivalent to the source language and both culturally and linguistically appropriate to the target country (Conway et al., 2010). Assessing the translatability of a measure during its development can therefore reduce problems encountered in future translations and avoid incorrect interpretations of items.

Lastly, this chapter explores postal and online recruitment strategies which may inform future studies using the eHIQ. There is no conclusive way of estimating the response rate for online questionnaires, particularly for open recruitment where a questionnaire link is made publicly available with unrestricted access. Previous suggestions or attempts to calculate response rates with open recruitment have been made by using the number of visitors to the study website as a denominator, by calculating the number of people who viewed the study advertisement on a website, by recording the number of clicks on a research invitation advertisement and by specifying the number of people who register for a study (Koo and Skinner, 2005, Graham et al., 2008). Other recommendations advocate using the number of people who access, but do not complete the questionnaire as a denominator (Eysenbach and Wyatt, 2002).

7.1.1 Objectives

A number of primary and secondary objectives of Stage Four are outlined below.

Objectives are not listed in sequence of execution, but in order of priority.

7.1.1.1 Primary

- To finalise the eHIQ sub-scales
- To confirm the convergent validity and external reliability of the eHIQ sub-scales

7.1.1.2 Secondary

- To refine the electronic design and format of the eHIQ
- To test the linguistic and cultural translatability of the eHIQ instructions, items and response options
- To explore the feasibility of recruiting participants to complete the eHIQ using a number of recruitment strategies

7.2 Research design

Having reduced the number of items in previous stages, a translatability assessment was carried out to identify any wording which would cause difficulty during translations into another language. The questionnaire was subsequently administered online. To ensure items were suitable for use across a range of patient and carer groups, the sampling framework aimed to reflect a breadth of health conditions and, where possible, was stratified for age and gender. Respondents were asked to complete the questionnaire on two occasions with a two week interval to test external reliability. Appropriate reference measures were included in the online survey to examine convergent validity.

Mixed modes of recruitment consisted of distributing postal research invitations and circulating an electronic link to the questionnaire. The electronic questionnaire link was circulated using email distribution lists and postings made publicly available on the internet.

7.3 Methodology

7.3.1 Translatability assessment

A translatability assessment of the 48 items was undertaken in collaboration with experts in the field (Pharma Quest Ltd). The assessment was managed by Pharma Quest and aimed to identify and amend wording which may prove difficult to accurately translate into another language. Items and instructions were assessed for cross-cultural and linguistic ease of translation into a number of carefully selected languages. Several stages were undertaken and are discussed in further detail below.

7.3.1.1 Concept elaboration

A concept elaboration document was produced which aimed to clarify any ambiguities or nuances within the questionnaire. To do this, Pharma Quest defined and expanded on the eHIQ instructions and individual items within a tabulated summary. The tabulated summary was sent to LK to be amended, verified and further expanded where necessary. The final document was then returned to Pharma Quest.

7.3.1.2 Translation assessment

Following advice provided by Pharma Quest, a combination of languages and countries were chosen by the research team (LK and supervisors CJ and SZ) which reflected a range of cultural and linguistic perspectives. The questionnaire was reviewed by the following selected language (and country) combinations:

- Arabic (Tunisia),
- French (France),
- German (Germany),
- Japanese (Japan),
- Simplified Chinese (China) and,
- Spanish (USA)

A copy of the questionnaire and the concept elaboration document were sent to medical translators living in each selected country. Translators were asked to provide feedback on whether the instructions and items could be easily translated into their native language from a cultural and linguistic perspective. Where translators thought a direct translation would not be appropriate for a specific country or language, they were asked to suggest more appropriate wording which conveyed the original meaning as defined in the concept elaboration document.

7.3.1.3 Teleconference and final report

Comments from the medical translators consulted were reviewed, summarised by Pharma Quest and sent to LK. A teleconference was held between representatives of Pharma Quest and LK to discuss the translator's comments and recommendations. Suitable adjustments to wording were discussed and a final report documenting the translator's comments, Pharma Quests' recommendation and LK's comments was produced.

7.3.1.4 Final amendments

LK, SZ and CJ discussed comments made by the translators and recommendations provided by Pharma Quest. A consensus was reached regarding appropriate amendments to wording within the questionnaire.

7.3.2 Establishing psychometric properties and recruitment feasibility

7.3.2.1 Participants

Men and women who had access to the internet, were aged 18 years or more, lived in the UK and were from one of the following population groups were included:

- Patients diagnosed with asthma
- Smokers referred to NHS smoking services within the last year
- People with multiple sclerosis (MS)
- Carers of people with MS
- People interested in knowing more about healthy eating
- People with motor neurone disease (MND)
- Cancer survivors
- Women who are experiencing the menopause (Pre/Peri/Post-menopausal)
- Carers (generic)

7.3.3 Procedure

7.3.3.1 Recruitment

A recruitment strategy was designed to achieve an appropriate sample size for statistical analysis (see section 7.3.7.2) and to evaluate the feasibility of recruiting participants via postal and online recruitment modes. Recruitment took place in two phases: Phase 1 recruited participants through mailing out hard copy research invitation packs and Phase 2 recruited participants via the internet. Data were also obtained from a website evaluation carried out concurrently with this study by a senior researcher within the

Health Experiences Research Group (HERG), University of Oxford. The purpose of incorporating this data set into Stage Four was to increase the sample size. The recruitment strategy used for this secondary analysis dataset was therefore not analysed in detail and participants were not asked to complete a second questionnaire for test-retest reliability. For ease of interpretation this secondary data is referred to as ‘Phase 3’ throughout this chapter.

7.3.3.2 Phase 1: Mail out recruitment

People who would like to stop smoking and people diagnosed with asthma were invited to take part in this research through the Oxfordshire Primary Care Trust. A cover letter (see Appendix 7.1) and information about the research (see Appendix 7.2) were distributed to General Practitioner (GP) practices. GPs were asked to recruit participants to the study by mailing patients within their practice. For ease of managing recruitment materials and to reduce the possibility for the same person being selected to complete a questionnaire for both health conditions (i.e. asthma and smoking cessation), each practice was asked to mail research invitations to one condition group only.

Staff within six participating GP practices were asked to conduct an electronic database search using predefined eligibility criteria (criteria outlined in section 7.3.7.2). Although harm to asthma and smoking cessation participants through participating in the study was considered low risk, identified patients were reviewed by a GP within the practice to exclude patients who may be adversely affected.

Carers of people with MS and people with MS were mailed through a membership list held by the Birmingham branch of the MS Society. Recruitment materials were provided to the branch and address labels were attached by volunteer workers within the branch (see Table 7-1).

7.3.3.3 Phase 2: Electronic recruitment

Recruitment took place using a variety of online sources. Direct recruitment took place through emailing potential participants through condition specific organisation, through educational organisations and through asking colleagues to forward the questionnaire link to people who they thought may be interested in taking part. Asking people to forward on emails (or chain referrals) have been found to be an inexpensive and effective method of recruiting individuals in a short period of time (Rock, 2010). Failed delivery emails were recorded where possible.

Open recruitment took place through advertising the study on relevant health-related websites and social networking groups. These websites included health blogs, online discussion forums, social networking sites (Facebook and Twitter) and news pages. Care was taken to advertise the study on websites which had not been previously used for recruitment in Stage Three to increase the likelihood of achieving independent samples. All discussion group administrators were asked for permission to post the advertisement to add legitimacy to the request (see Table 7-1) (Eysenbach and Wyatt, 2002).

7.3.3.4 Phase 3: Website evaluation recruitment

Direct recruitment took place through emailing potential participants on an email list belonging to a National Cancer Research Network Consumer Liaison Group. Participants were also recruited through asking a variety of individuals and organisations to forward the questionnaire link to people who they thought may be interested in taking part.

Open recruitment took place through the DIPEx charity who advertised the questionnaire on their Twitter and Facebook accounts. Advertisements were also posted on relevant websites and discussion boards. See Table 7-1 for further information on recruitment sources.

Table 7-1 Stage Four Recruitment sources

Population	Direct recruitment	Open recruitment
Phase 1		
Smoking cessation	Oxfordshire PCT • Didcot Health Centre • The Health Centre, Bicester • Windrush Medical Practice	-
Asthma	Oxfordshire PCT • Jericho Health Centre • 28 Beaumont Street • Summertown Health Centre	-
People with MS	MS Society Birmingham branch	-
MS (carers)	MS Society Birmingham branch	-
Phase 2		
Healthy eating	Oxford Brookes University • Student mail list (approximately 325 students) Email circulation lists • Snowballing (colleagues) • HERG research participant panel list (21 members)	Facebook
People with MS	MS Society branches • Furness branch (100 members) • Newcastle and Gateshead branch (50 members) • Isle of Man • Berwick and Eastern Borders (Newsletter) Oxford MS Therapy Centre (Mail list)	Facebook Twitter
People with MND	Motor Neurone Disease Association • e-Communications mail list (250 members)	-
Menopause	-	Menopause Matters • Discussion forum
Generic carers	-	Patient Talk • Blog
Phase 3		
Cancer survivors	National Cancer Research Network Consumer Liaison Group (Mailing list) Emails distribution: (Thames Valley Cancer Network, Department of Primary Care Health Sciences, Cancer charities)	Macmillan • Community noticeboard • Cancer Voices (database of several thousand people who wish to be involved in PPI opportunities)

7.3.4 Materials

7.3.4.1 Phase 1: Mail out recruitment

Participants recruited from GP practices were mailed a research invitation pack which included a cover letter (see Appendices 7.3 and 7.4) and a participant information sheet (see Appendices 7.5 and 7.6). A reminder letter was mailed to each patient ten days after the initial invitation pack had been sent (see Appendices 7.7 and 7.8). A member of staff within participating GP practices attached address labels to the sealed questionnaire pack ensuring the anonymity of potential recruits.

Members on the Birmingham MS Society branch list were mailed a hard copy research invitation pack consisting of a cover letter (See Appendix 7.9), and a participant information sheet for a person with MS (see Appendix 7.10) and, if applicable, a participant information sheet for their carer (see Appendix 7.11). A reminder letter was mailed to each member ten days after the initial invitation pack had been sent (see Appendix 7.12).

All participants were provided with a relatively short web address which directed them to the University of Oxford's *Health Services Research Unit* (HSRU) website. The specified page on the HSRU website contained the participant information again and a questionnaire link which a participant could click on if they wished to take part in the study. The questionnaire link was placed at the top of the webpage, within the participant information text and also on the bottom of the webpage. Two weeks after completing the questionnaire, participants were sent an email asking them to complete the questionnaire again for the test-retest questionnaire (see Appendix 7.13).

7.3.4.2 Phase 2: Electronic recruitment

Participants were either emailed a research invitation (see Appendix 7.14) directly to their inbox or invited to take part in the study through an advertised post on a relevant website (see Appendix 7.15). All potential participants were given access to an electronic link which they could click on or copy and paste directly into their web browser. The electronic link lead directly to the front page of the survey and participants could view the participant information sheet on the second page of the survey (see Appendix 7.16). Participants were sent an email asking them to complete the questionnaire again following a period of two weeks (see Appendix 7.17).

7.3.4.3 Phase 3: Website evaluation recruitment

An invitation letter was emailed to members of a National Cancer Research Network Consumer Liaison Group mailing list asking participants to complete the eHIQ-Part 1 and given access to the website under evaluation for seven days. Reminder emails were sent to all participants who had completed eHIQ-Part 1 within seven days asking them to complete eHIQ-Part 2. Second reminders were sent to participants who had not completed eHIQ-Part 2 one week after the initial reminder had been sent.

7.3.5 The survey design

Qualtrics survey software program was used to design the electronic questionnaire (Qualtrics, 2005). The decision to use Qualtrics software rather than BOS software (used in Stage Three) reflected the increased control and manipulation possible with Qualtics software during the survey design process. While BOS software was functional, it became apparent that administering the questionnaire would benefit from some additional features which were available using Qualtrics. In addition to achieving a more user friendly and visually appealing questionnaire, refining and tailoring the questionnaire design can reduce measurement error as layout can reduce possibility of response bias and items being overlooked (Dillman, 2011). Improvements which were possible by changing the software are outlined in Table 7-2 and illustrated in Appendix 7.18. A further adjustment to aid ease of completion involved limiting the number of items appearing on one webpage so that the respondent was not required to ‘scroll down’ to access the ‘Continue’ button at the bottom of the page. The use of colour was also restrained in the text so that consistency and readability were prioritised. As with the survey distributed in Stage Three, contact information was provided on the bottom of each survey page.

Answering each question in the survey was not mandatory; however a prompt box appeared where participants had left an item unchecked. The survey software also had an inbuilt survey mailer which was used for the retest (Time 2) questionnaire. Participants’ email addresses and identifiers were imported into the software which distributed customised questionnaire links to the retest survey (Dobrow et al., 2008). This enabled automated reminders to be sent to non-responders and informed the linking of the Time 1 and Time 2 questionnaires.

7.3.6 Ethical approval and participant consent

Ethical approval was obtained for Phase 1 recruitment through the NHS Research Ethics Committee (REC reference number: 12/SW/0209). Ethical approval for Phase 2 recruitment was obtained through the Medical Sciences Inter Divisional Research Ethics Committee (MSD-IDREC-C1-2013-063). Before completing the questionnaire, participants were asked to provide their consent to take part in the study by clicking on consent boxes provided (see Appendices 7.19-7.20).

Table 7-2 Stage Four survey design improvements

Design issue	Bristol Online Survey software	Improved function using Qualtrics software
Standardised format	Grid format questions varied in width according to participants' computer monitor size.	Grid format questions with a fixed width survey template (See Figure 1, Appendix 7.18). This made the format of the questions as standardised as possible regardless of the respondent's computer monitor size.
Monitoring completion progress	A progress bar was requested by a number of participants during the cognitive debrief interviews. This function is not available on BOS software. It is possible that the lack of this feature this can contribute to participant dropout.	Inclusion of a progress bar on the bottom of each page (see Figure 2, Appendix 7.18).
Navigation	No facility to adjust submit buttons (for example, increase size/move position).	Ability to increase the appearance and position of the submit buttons.
	No facility available to insert reminder text boxes on webpages.	Ability to include text reference boxes to highlight important information or instructions to participants.
Ballot stuffing	Unable to record IP address or ascertain if a respondent completed the questionnaire twice.	Facility to record IP address and to prevent multiple answers being submitted from the same computer.
Condensing information	Limited use of computer coding (HTML and CSS) to adjust appearance of the survey and condense large amounts of information.	Ability to insert computer coding to hide text. This allowed large amounts of text in the participant information section to be condensed. Participants are able to click on a question in order to get the answer in drop down format.
Managing retest responses	Respondent required to enter identification variable in order to match responses on first and second questionnaire completion.	Inbuilt survey mailer facilitates easy linking of questionnaires. Without relying on information from participant to re-enter data.

7.3.7 Data collection

On providing consent to take part in the study, participants were asked to provide their email address, complete basic demographic questions, the eHIQ-Part 1 and a reference item to assess convergent validity. Participants then spent 10-15 minutes browsing a specified health-related website, following which, participants were asked to complete the eHIQ-Part 2, reference scale items to assess convergent validity, further demographic questions and items relating to internet use. See Appendix 7.21 for Stage Four eHIQ items and numbering.

Seven of the eight internet frequency items related to general internet use (Blank, 2010). Mean scores of the seven items were calculated providing an insight into how familiar participants were with using the internet in their everyday life. The seven items asked how frequently participants used tools such as email accounts, internet blogs, internet discussion boards, chat rooms, media tools and social networking. For the purposes of calculating scores, those selecting the response option ‘Don’t know’ were excluded from the analysis. Scores could range from 1 (most frequent use of internet tools) and 7 (least frequent use of internet tools). One adhoc item was included about frequency of use of health-related websites.

Each survey administered also contained condition specific items. Condition specific items were not part of the eHIQ, but used for descriptive analysis of the sample. An open ended question was included at the end of the survey so that participants could provide any comments if they so wished. See Table 7-3 for items included in each survey. IP addresses were recorded through the survey software.

7.3.7.1 Reference measures

The reference measure included to test convergent validity with the eHIQ-Part 1 consisted of a single item from the Health Information National Trends Survey (HINTS) (Nelson et al., 2004). Single items can be considered useful in measuring a concept (Jenkinson et al., 1994) and it was hypothesised that a single item on trust would be moderately correlated with the eHIQ-Part 1 sub-scales. The single item was taken from the HINTS questions database of questions asked in 2007: *In general, how much would you trust information about health or medical topics on the internet? (A lot/Some/Little/Not at all)*.

Findings from the structured review (Chapter 3) informed the selection of the reference measure included to test convergent validity with the eHIQ-Part 2. One sub-scale of the *Web Trust Questionnaire* (Access to quality information, eight items) (Harris et al., 2011) was included as it was believed that this scale should have moderate correlations with all sub-scales within the eHIQ-Part 2. Responses are given on the same five point scale as the eHIQ ranging from 'Strongly disagree' to 'Strongly agree'. The *Access to quality information* sub-scale has demonstrated good internal consistency (Cronbach's alpha =0.85) and describes ease of use and understanding of website content. Items included in the sub-scale are listed in Box 7.1. Note that two of the eight items (Items 1 and 3) are the same as two items included in the eHIQ-Part 2 questionnaire. Analyses therefore also involved comparing the relationship of the eHIQ sub-scales with an adjusted six item scale which excluded the two overlapping items. Scores for the *Access to Quality Information* sub-scale were obtained by calculating the mean item score. For the adjusted six item scale, the mean score of the included six items were calculated.

Box 7.1 Access to quality information sub-scale

Eight items:

1. **The language on the site made it easy to understand.**
2. The site helped me understand the issue better.
3. **The site was easy to use.**
4. The site told me most of what I needed to know.
5. The layout was consistent with other sites.
6. The advice appeared to be prepared by an expert.
7. The advice seemed to be offered in my best interests.
8. The advice came from a knowledgeable source.

(Harris et al., 2011)

*Items appearing in bold are also present in eHIQ-Part 2

Table 7-3 Stage Four Items administered by population group

Population group	Item content
All participants	
eHIQ-Part 1	Item bank 1: General attitudes towards using the internet for health information (14 items)
eHIQ-Part 2	Item bank 2: The impact of using a specific health-related website (34 items)
Demographics	Age, Sex, Education, Ethnicity, Employment status
Internet use	Frequency items adapted adapted from the Oxford Internet Institute Survey (Blank, 2010)
eHIQ-Part 1 reference measure	HINTS survey question (Nelson et al., 2004)
eHIQ-Part 2 reference measure	Access to quality information scale (Harris et al., 2011)
Population specific	
Asthma	Asthma diagnosis, Length of time since diagnosis, Severity
Smoking cessation	Age when started smoking, Current feelings on smoking
People with MS and MND	Age when diagnosed
Carers (MS and Generic)	Length of time being a carer
People who have survived cancer	Cancer type
Menopause	Menstrual status
Healthy eating	World Health Organistaion ‘Diet’ core questions (WHO, 2008)

7.3.7.2 Sampling and sample size

As outlined in the previous chapter, estimates suggest that meaningful psychometric tests require at least three times as many respondents as items (Kline, 2000). The longest part of the eHIQ (eHIQ-Part 2) contained 34 items, therefore a suitable sample size was estimated at just over 100 participants. This figure served as a minimum recruitment target, however as other estimates encourage a larger size (Hinkin, 1998).

7.3.7.3 Phase 1: Mail out recruitment

Due to the demographic information held on GP practice databases, a pre-defined sampling criteria was set (see Table 7-4). Each GP practice was asked to select 90 particpants from their database stratified by age and gender (see Appendices 7.22-23) for instructions to practices). The total number of patients GP practices were asked to contact was therefore 540. The large number of patients contacted reflected the relatively low response rates experienced in previous studies which have administered questionnaires through Primary Care, for example, the national GP Survey received a response rate of 35% during July 2012- April 2013 (NHS_England, 2013a).

The MS Society Birmingham branch had a total of 235 members on their mailing list. In the interests of achieving the maximum sample size possible, all members were contacted. Overall, it was planned that a total of 775 people would therefore receive a mailed a hard copy research invitation pack and a reminder letter.

Table 7-4 Stage Four Sampling criteria for GP practices

Population	Sampling criteria
Smoking cessation	≥ 18years Registered at the GP practice in question Categorised under one or more of the following Read[1] codes during the past year: o 8CAL- Smoking cessation advice given o 8H7i- Referral to smoking cessation advisor o 9N2k -Seen by stop smoking advisor o 8HTK- Referral to stop smoking clinic
Asthma	≥ 18years Diagnosed with asthma Registered at the GP practice in question

7.3.7.4 Phase 2: Electronic recruitment

Mailing research invitation packs in Phase 1 enabled contacting a pre-defined population and the calculation of a response rate. It was not possible to calculate an accurate response rate for Phase 2 due to an open recruitment process where the value of the denominator was uncertain (i.e. how many potential participants received or viewed the electronic research invitation).

It was possible to estimate how many emails were sent directly to potential participants email inboxes for two conditions (i.e. the number of people on a mailing list or the number of subscribers to an eCommunication list). This gave an indication of the likely number of individuals who received the electronic research invitation.

---

[1] Read Codes are a coded thesaurus of clinical terms which are used to classify patient records in primary and secondary care databases in the NHS.

In the absence of calculating accurate response rates for electronic recruitment, the IP addresses collected by the survey software were analysed to give some indication of the success of the recruitment strategy. In this instance the total number of people accessing the questionnaire (i.e. the number of IP addresses viewing the first page of the questionnaire) was considered the total number of potential participants sampled. Completion rates (those who completed the questionnaire divided by those who had consented to take part) were also calculated for both Phase 1 and Phase 2 participants.

7.4 Data analysis

Data for each population group were exported from Qualtric's survey software to SPSS files. Files were integrated to construct a master data set which contained data from all the condition groups collected. Duplicate IP addresses were identified within the dataset and blank duplicate entries were removed. Blank entries were likely to have occurred when a participant clicked on the questionnaire link to read the opening page, choosing to come back and complete the questionnaire at a later time. Items were recoded where appropriate so that responses were scored in the same direction where 1= low perceived value of the internet (or website) and 5= high perceived value of the internet (or website).

7.4.1 Finalising eHIQ sub-scales and scale scores

The internal consistency of the sub-scales which were identified in Stage Three were examined (Cronbach's alpha > 0.7) and individual item distributions were examined across response options for floor and ceiling effects. Where sub-scales did not display good internal consistency, an exploratory factor analysis was carried out (see Chapter 6 for methodology) to identify ill-fitting items and finalise optimal sub-scales.

A scoring system was established for sub-scales so that scores could be used in subsequent analysis. A simple unweighted summation scoring algorithm (DeVellis, 2006) was produced for each sub-scale which transformed to a 0-100 metric.

7.4.2 Statistical analysis

Descriptive statistics were used to describe the sample and the sub-scales score distributions. Scores were also examined for floor and ceiling effects. Relationships between the scales and a range of population characteristics are examined to identify potential covariate factors (gender, age and recruitment phase) related to population characteristics.

Comparisons between categorical and continuous variables were made using t-tests or ANOVAs and Pearson's correlation coefficients (r) were used for comparisons between continuous variables (Williams, 2009, Norman and Streiner, 2000). To carry out ANOVAs when comparing continuous variables with more than two groups, Levene's test for equality of variances was used to assume equality (or homogeneity) of variance among each sub-group. A significant ($p<0.05$) Levene's test indicates sub-groups are of different sizes, hence this test would be inappropriate (Norman and Streiner, 2000). On meeting the requirements for equality of variance, the F statistic was calculated. This takes into account the variability among the means and the standard errors of the means. Degrees of freedom (df) were reported in the SPSS output for the numerator (number of groups minus one) and the denominator (total sample minus the number of groups being compared). Post-hoc tests were used on significant ($p<0.05$) results in order to tell which groups were significantly different. The proportion of variance of association was calculated by dividing the sum of squares between groups by the total sum of square (Norman and Streiner, 2000).

Data were inspected for normal distribution. Where data appeared to have a non-normal distribution, non-parametric tests were substituted as follows: Spearman's correlations in place of Pearson's, Mann Whitney U for t-tests and Krustal-Wallis for ANOVAs (Williams, 2009).

To establish reliability of the measure the test-retest procedure was used. The level of agreement between scores from the two occasions was assessed using the intra-class correlation coefficient (ICC). This is a standard reliability test for continuous data and coefficients are considered psychometrically satisfactory above 0.70 and particularly good if they are above 0.80 (Nunnally and Bernstein, 1994).

7.5 Results

This section begins with a description of the recommendations made following the Translatability assessment and a description of the resulting questionnaire amendments. A description of the study sample is then reported in detail and recruitment issues arising during each phase of recruitment are discussed. The statistical analysis of the eHIQ data is then documented including the steps taken to finalise the sub-scales within the questionnaire, evidence of convergent validity, descriptions of sub-scale score distributions and finally, evidence of test-retest reliability.

7.5.1 Translatability Assessment

Several minor changes to the instructions and item wording within the eHIQ were thought to be necessary following the translatability assessment. Changes were suggested to avoid inconsistency with wording, to ensure all wording was culturally and linguistically relevant to the selected countries and to avoid ambiguous wording or poor grammar which may jeopardise future translations. For the full list of suggestions contained within the translatability report provided by PharmaQuest see Appendix 7.24.

Seven minor modifications were made to improve grammar and consistency with wording. Two items were amended to improve their cultural or linguistic suitability (see Table 7-5) and six items were amended to avoid ambiguous wording (see Table 7-6 for examples). Three further items were identified as potentially being problematic in future translations. These items were:

- Q19) I feel I have a sense of solidarity with other people using the website.

- Q21) I can identify with other people using the website.
- Q26) I feel I have a lot in common with other people using the website.

While the above items had similar, yet subtly different meanings in the English language, translators were concerned some languages may not have the capacity to express those differences (for example, due to fewer adjectives or forms of expression in the target language). Despite the translators concerns however, it was decided to retain the items in the interests of capturing the intended concept fully in English.

7.5.2 Data preparation and analysis

In cases where there was more than one questionnaire completed under the same IP address further investigation was carried out to determine if more than one response had been submitted by the same respondent. In two cases it was considered reasonable to have more than one questionnaire completed under the same IP address. In the first case a questionnaire for Carers of people with MS and a questionnaire for People with MS shared a common IP address. The two entries had different email addresses and demographic details. In this instance it was reasonable to assume the questionnaire had been completed by a person with MS and by their carer both using the same internet access point. In the second case, five questionnaires were submitted under the same IP address. Further investigation of the five entries established they all had a different email address, however, shared a common email domain (i.e. their place of work). The five participants also had different responses to demographic questions. In this instance, the five questionnaires were retained. No other completed questionnaires shared a common IP address.

Table 7-5 Stage Four Linguistic and cultural translatability concerns

Item	**Is a direct translation linguistically possible and appropriate in the target language?**	**Is a direct translation culturally appropriate for the target country?**	**PharmaQuest Comments on Linguists' Suggestions**
I would use the internet if I needed help to make a decision about my health (for example, whether I should seek medical advice, take medication or seek alternative therapies).	**French (France):** The concept of 'alternative therapies' … may need elaborating. **Japanese (Japan):** A direct translation of "medical advice" may imply "advice from Western medicine professionals". **Simplified Chinese (China):** Suggests alternative wording to sound more idiomatic in Chinese.	**Japanese (Japan):** It may be necessary to elaborate on what "alternative therapies" refer to, because the perception may differ depending on the county/culture. For example, in Japan acupuncture may be considered mainstream by some people.	**PharmaQuest:** Suggested wording: I would use the internet if I needed help to make a decision about my health (for example, whether I should see a doctor, take medication or seek other types of treatment).
This website encourages me to take steps that could be beneficial to my health.	**Arabic (Tunisia):** Although there's a note to translators explaining that "to take steps" should not be understood in its literal meaning, a mistranslation is always possible. **Simplified Chinese (China):** "Take actions" is more commonly used in this context than "take steps" in Chinese.	Yes	**PharmaQuest:** It is possible that 'take steps' will cause confusion during the translation process. 'Take actions' is still an active, positive concept that accurately conveys the meaning of 'take steps'.

Table 7-6 Stage Four Examples of ambiguous wording

Item	Ambiguity	PharmaQuest Comments on Linguists' Suggestions
The internet is useful if you don't want to tell people in your everyday life how you feel.	**PharmaQuest:** 'People in your everyday life' is quite a vague concept that the translators had an issue with. It could also be beneficial and clearer to change 'people in your everyday life' to 'the people around you'.	The internet is useful if you don't want to tell people around you (for example, your family or people at work) how you feel.
I would not want someone I care for to see this website.	**PharmaQuest:** Suggest changing 'someone I care for' to 'someone I care about'.	I would not want someone I care about to see the website.

7.5.3 Characteristics

The two part eHIQ was completed online by 170 participants. Phase 1 recruited 47 participants, Phase 2 recruited 96 participants and Phase 3 recruited 27 participants. Of the 170 respondents, 59 were men and 107 were women with four respondents not disclosing their sex. Women outnumbered men in all population groups with the exception of MND (see Table 7-7). The most common forms of cancer in Phase 3 were breast cancer (n=11, 40.7%) and head and neck cancers (n=3, 11.1%).

The mean age of the Phase 1 and 2 samples was 50.4 years (SD 13.02) and the modal age band for Phase 3 was 51-75 years old[1]. All participants had completed secondary level

---

[1] Phase 1 and 2 participants were asked to enter their age while participants in the Phase 3 were asked to select the age band they belonged to.

education with 128 (76.6%) participants having completed third level education. The most common employment status was 'employed' (n=75, 44.6%) followed by 'retired' at retirement age or due to ill health, (n=46, 28.6%). The vast majority of the sample classified themselves as 'White' (n=160, 95.8%). See Table 7-8 for further breakdown of demographic characteristics.

The mean age of diagnosis for people with MND was 55.1 (SD 7.1) while the mean age of diagnosis for people with MS was 37.3 (SD 9.5). The mean number of years since diagnosis for people with asthma was 15.9 (SD 16.8). The mean number of years since cancer diagnosis for cancer survivors was 4.6 (SD 1.9). See Table 7-9 for further condition specific details.

Table 7-7 Stage Four Condition groups according to recruitment phase

Sample	Male	Female	Total
Phase 1			
Smoking cessation	0	3	3
Asthma	8	9	17
Multiple Sclerosis	4	16	20
Multiple Sclerosis (carers)	2	2	4
Phase 2			
Healthy eating	15	33	48
Multiple Sclerosis	1	5	6
Motor Neurone Disease	20	12	32
Menopause	0	7	7
Generic carers	1	1	2
Phase 3			
Cancer survivors	8	19	27
Total	59	107	166

Missing n=4

Table 7-8 Stage Four Participant characteristics

Participant	**n (%)**			
	Phase 1	**Phase 2**	**Phase 3**	**Total**
Sex				
Male	14 (31.8)	37 (38.9)	8 (29.6)	59 (35.5)
Female	30 (68.2)	58 (61.1)	19 (70.4)	107 (64.5)
Missing	3(-)	1 (-)		
Age				
Mean	54.8 years (SD 14.1)	48.2 years (SD 11.9)	-	50.4 years (SD13.0)
Range	57 years	51 years	-	59 years
Mode*	-	-	51-75 years (n=17)	-
Education completed				
Secondary school	15 (31.9)	16 (16.8)	8 (32.0)	39 (23.4)
Third level	32 (68.1)	79 (83.2)	17 (68.0)	128 (76.6)
Missing	(-)	1 (-)	2 (-)	
Employment status				
Employed	19 (40.4)	45 (47.4)	11 (42.3)	75 (44.6)
Unemployed	1 (2.1)	3 (3.2)	0 (0)	4 (2.4)
Student	2 (4.3)	5 (5.3)	0 (0)	7 (4.2)
Not working (ill health)	3(6.4)	14 (14.7)	4 (15.4)	21 (12.5)
Home duties	2 (4.3)	6 (6.3)	2 (7.7)	10 (6.0)
Retired (Retirement age)	11 (23.4)	8 (8.4)	3 (11.5)	22 (13.1)
Retired (Ill health)	9 (19.1)	14 (14.7)	3 (11.5)	26 (15.5)
Other	0 (0)	0 (0)	3 (11.5)	3 (1.8)
Missing	(-)	1 (-)	1 (-)	
Ethnicity				
White	46 (97.9)	90 (94.7)	24 (96.0)	160 (95.8)
Black Caribbean	0 (0)	1 (1.1)	0 (0)	1 (0.6)
Asian	0 (0)	2 (2.1)	0 (0)	2 (1.2)
Mixed race	0 (0)	0 (0)	1 (4.0)	1 (0.6)
Other	1 (4.3)	2 (2.1)	0 (0)	3 (1.8)
Missing	(-)	1 (-)	2 (-)	3 (-)

*Mode only reported for Phase 3 as Phase 1 and 2 participants were asked to enter their age while participants in the Phase 3 were asked to select the age band they belonged to.

Table 7-9 Stage Four Condition specific participant characteristics

Sub-sample	Frequency/ Mean (SD)
Smoking cessation (n=6)	
Would like to stop smoking someday	1
Going to stop smoking in the next month or less	1
Stopped smoking over six months ago	2
Missing	2
Healthy eating (n= 49)	
Three or more servings in a day:	
Fruit	16
Vegetables	29
Asthma (n=17)	
Severity	
Mild	12
Moderate	5
Years since diagnosis	15.9 (SD 16.8)
MND (n=31)	
Age when diagnosed	55.1 (SD 7.7)
MS (n=32)	
Age when diagnosed	37.3 (SD 9.5)
Menopause (n= 7)	
Perimenopausal /transition	2
Postmenopausal	4
Other	1
Cancer (n=27)	
Years since diagnosis	4.6 (SD 1.9)

7.5.3.1 Internet use

Of those who responded, 143 (85.1%) confirmed they had visited a health website before. Mean scores were calculated for each recruitment phase to ascertain if there were differences in internet use according to mode of recruitment (see Table 7-10). The seven item scale was checked for internal consistency (Cronbach's alpha= 0.80) and the mean internet use score for the sample was 4.19 (SD 0.98), where 1=most frequent use of the internet and 7= least frequent use of the internet. Those recruited in the mail out research invitation (Phase 1) appeared to use internet tools less frequently than those recruited online (Phase 2 and Phase 3).

A one way ANOVA confirmed significant differences in frequency of internet use among the three recruitment phases [$F_{(2,155)}$ =7.776, p=0.001]. Post-hoc comparisons using Tukey's statistic indicated the significant differences were between Phase 1 participants (Mean=4.64, SD 0.86) and Phase 2 participants (Mean=3.97, SD 0.97).

Table 7-10 Stage Four Mean internet frequency

Recruitment phase	n	Mean* (SD)
Phase 1	45	4.64 (0.86)
Phase 2	88	3.97 (0.97)
Phase 3	25	4.13 (0.94)
Total	158	4.19 (0.98)

*Calculated using 7 items from the Oxford Internet Survey where 1=most frequent use of the internet and 7= least frequent use of the internet

7.5.4 Recruitment

Recruitment levels for all recruitment phases was disappointing. Phase 1 and 2 recruitment is discussed in further detail below.

7.5.4.1 Phase 1

In total, 755 research invitation packs and reminder letters were mailed to potential participants through primary care or the MS Society Birmingham branch. This number was lower than predicted as one GP practice could only mail research invitation packs to 70 of the requested 90 asthma patients. This achieved an overall response of 47 (6.1%) participants from the 755 invites. Of the 250 people with asthma who were asked to take part in this study, 17 (6.8%) completed the questionnaire. Of the 270 people identified from GP practice lists who would like to stop smoking, six (2.2%) completed the questionnaire. Using the MS Society Birmingham branch mailing list, 235 members were contacted yielding a response of 20 people with MS and four carers of people with MS. The response rates for the MS groups are unknown as it was impossible to establish how many of those listed were either people with MS or carers of people with MS. Four reminder letters were returned to sender from the MS mail out as the administrator did not place an address label on the envelope. Due to the extremely poor response rate, the Birmingham MS Society also emailed members who had registered email addresses with the organisation. Therefore, 132 members were also sent an email reminder research invitation to encourage uptake.

During Phase 1 recruitment, 14 people telephoned and three people emailed LK to explain why they could not take part in the study. Reasons included not having computer access;

the addressee no longer living at the address or not having the specified condition. Three people got in contact as they had trouble accessing the online questionnaire. All three participants were able to access the questionnaire with help from LK. See Table 7-11 for further detail.

For those who could not access the questionnaire, no obvious technical fault seemed to exist, however, participants were required to enter a web URL into their browser and click on a link embedded on the University departmental webpage. Although efforts were made to make this process as simple as possible through detailed instructions, using a simple web URL and providing a number of electronic links to the questionnaire on the departmental page, this proved difficult for some participants.

Table 7-11 Stage Four Participant feedback

Sample	**Mode of contact (Frequency)**		
	Phone	**Email**	**Total**
Asthma			
Does not have condition (or very mild)	1	2	3
Not at address (return to sender)	1	-	1
No computer access	3	-	3
Needed help to access survey	-	1	1
Smoking cessation			
No computer access	2	-	2
Needed help to access survey	1	-	1
Multiple Sclerosis			
Does not have condition	2	1	3
Needed help to access survey	1	-	1
Unknown			
No computer access	5	-	5
Total	16	4	20

7.5.4.2 Phase 2

Phase 1 recruitment was disappointing and further highlighted the need to make accessing the questionnaire as user friendly as possible (i.e. accessing the online URL). One way to do this was to decrease the number of links a participant would need to navigate in order to begin completing the questionnaire. To ease questionnaire access in Phase 2, the study web link led directly to the front page of the survey (i.e. participants by-passed the HSRU website). Participant information text was then included on the second page of the questionnaire. Participants could click or hover over participant information questions to read the answer. On reading the further information, participants could click on 'Continue' to complete the consent form if they wished to participate.

It was possible to estimate the response rate for two conditions in Phase 2. For MND, 250 people were sent an email (or e-communication) from the MND Association. Thirty two people with MND completed the questionnaire achieving a response rate of 12.8%. The MS Society branches involved in the study emailed a research invitation to 150 people and achieved a response from six participants (4%). These response rates are estimates and do not account for participants forwarding on research invitations to people they think may be interested in taking part in the study. The email research invitations were sent by the administrators in each organisation and no information regarding failed emails was received from the administrators by LK.

After excluding duplicate IP addresses, 44 people gave consent, but did not complete the full questionnaire, 96 people gave consent and completed the questionnaire and 115 people clicked on the questionnaire link to access the front page, but did not continue to provide consent or complete any questions. The total number of people accessing the questionnaire was therefore approximately 225. Taking into account the 96 people who completed the questionnaire compared to the number of people viewing the front page of the study, a response rate of 42.7% was calculated.

7.5.4.3 Completion rates

Completion rates were calculated to reflect the number of participants who completed the questionnaire having given consent to take part in the study. In Phase 1, 47 (77%) of consenting participants completed the questionnaire while 96 (68.6%) of participants in Phase 2 completed the questionnaire (see Table 7-12). A higher rate of completion was expected in Phase 1 as the participant was invited to take part via a letter posted to their address. Therefore they had made the decision to access the questionnaire online in the first instance. Those logging onto their computer to access the online questionnaire were arguably more committed or invested in the study than those who accessed while online.

Table 7-12Stage Four Completion rates

Sample	Complete (n)	Non-complete (n)	Completion rate (%)
Phase 1			
Smoking cessation	6	6	50.0
Asthma	17	2	89.5
Multiple Sclerosis	20	6	76.9
Multiple Sclerosis (carers)	4	0	100.0
Sub-total	47	14	77.0
Phase 2			
Healthy eating	49	19	72.1
Multiple Sclerosis	6	5	54.5
Motor Neurone Disease	32	15	68.1
Menopause	7	0	100.0
Generic carers	2	5	28.6
Sub-total	96	44	68.6
Total	143	58	71.1

7.5.5 Finalising the sub-scales (eHIQ- Part 1)

Using the Stage Four dataset, both the internal consistency of each sub-scale identified in Stage Three and the distributions across response options for each item were inspected. Two of the three sub-scales displayed good internal consistency with Cronbach's alpha values of 0.71 or higher. One sub-scale (Attitudes towards online health information) however indicated lower internal consistency than desired (Cronbach's alpha = 0.60). It was noted that this value would improve if Item 4 (The internet causes people to worry about their health) were removed. In order to find the optimal factor solution, all eHIQ-Part 1 items (i.e. including Item 4) were entered into an exploratory factor analysis to examine the questionnaire structure again more thoroughly.

7.5.5.1 Factor analysis (eHIQ-Part 1)

The 14 eHIQ-Part 1 items were assessed for their suitability to factor analysis. The Kaiser-Meyer-Olkin (KMO) value exceeded the recommended value of 0.6 (KMO= 0.85) and the Bartlett's Test of Sphericity reached statistical significance ($p<0.01$) ensuring there was correlation between the items and thus supporting the use of factor analysis.

Four factors with Eigenvalues above one were initially extracted. These factors accounted for 64.19% of the variance. Closer inspection of the factor loadings indicated Item 4 did not group together with other similar items on a given factor. Item 4 also had poor distribution across response options with no participant selecting the 'Strongly disagree' category. Participants may have felt that this item (The internet causes people to worry about their health) was too general or perhaps poorly worded as it is more likely that information on websites cause people to worry as opposed to 'the internet'. Due to uncertainties regarding item clarity, its meaning and its poor psychometric properties, item 4 was removed.

Following the removal of Item 4, all remaining items were entered into a further factor analysis. Three factors were extracted explaining 59.97% of the variance. The three factors were subjected to a Direct Oblimin rotation and the Structure matrix and the Pattern matrix output were used to interpret the factor structure. As with the previous Stage Three analysis, the Structure Matrix offered primary guidance, but it was noted that the Pattern Matrix loadings allocated the same items to each factor. The three factors were assessed for internal consistency however the two item sub-scale with Item 1 (I know how to find relevant health information on the internet) and Item 3 (I know which health information on the internet is trustworthy) had a poor Cronbach's alpha value of 0.59.

Further examination of the two item sub-scale highlighted that these items were the only items within the questionnaire to begin with 'I know…'. It is possible participants answered these items in a different manner than the other items within the questionnaire. For example, the stem may be suited to a dichotomous response option format. It was also thought that issues around trust and relevance may have been picked up in other items in the remaining sub-scales. As it is preferable to have three or more items within a sub-scale and a higher Cronbach's alpha than 0.59, the two items were removed. Following their removal, factor analysis demonstrated the two remaining sub-scales accounting for 56.58% of the variance and the Structure matrix (see Table 7-13) and the Pattern matrix (see Table 7-14) output demonstrated stable sub-scales.

The two sub-scales retained were entitled: 1.1) *Attitudes towards online health information* and 1.2) *Attitudes towards sharing health experiences online* (see Table 7-15). Both sub-scales demonstrated good internal consistency and were conceptually relevant to overall attitudes towards online health information.

Table 7-13 Stage Four Structure matrix (eHIQ-Part 1)

Item	**Factor**	
	1	**2**
13. The internet is a good way of finding other people who are facing health-related decisions I may also face.	0.84	0.45
9. The internet is a good way of finding other people who are experiencing similar health problems.	0.84	0.31
10. It can be helpful to see other people's health-related experiences on the internet.	0.84	0.48
14. Looking at health websites reassures me that I am not alone with my health concerns.	0.77	0.49
11. The internet is useful if you don't want to tell people around you (for example, your family or people at work) how you feel.	0.67	0.38
12. It can be reassuring to know that I can access health-related websites at any time of the day or night.	0.62	0.59
7. I would use the internet if I needed help to make a decision about my health (for example, whether I should see a doctor, take medication or seek other types of treatment).	0.41	0.80
6. The internet can be useful to help people decide if their symptoms are important enough to go to see a doctor.	0.39	0.79
8. I would use the internet to check that the doctor is giving me appropriate advice.	0.37	0.71
5. The internet can help the public to know what it is like to live with a health problem.	0.39	0.65
2. The internet is a reliable resource to help me understand what a doctor tells me.	0.32	0.61

Extraction Method: Principal Component Analysis.

Table 7-14 Stage Four Pattern matrix (eHIQ-Part 1)

Item	Factor	
	1	2
9. The internet is a good way of finding other people who are experiencing similar health problems.	0.93	-0.18
13. The internet is a good way of finding other people who are facing health-related decisions I may also face.	0.83	0.02
10. It can be helpful to see other people's health-related experiences on the internet.	0.81	0.06
14. Looking at health websites reassures me that I am not alone with my health concerns.	0.71	0.11
11. The internet is useful if you don't want to tell people around you (for example, your family or people at work) how you feel.	0.65	0.04
12. It can be reassuring to know that I can access health-related websites at any time of the day or night.	0.43	0.36
6. The internet can be useful to help people decide if their symptoms are important enough to go to see a doctor.	-0.03	0.81
7. I would use the internet if I needed help to make a decision about my health (for example, whether I should see a doctor, take medication or seek other types of treatment).	-.013	0.81
8. I would use the internet to check that the doctor is giving me appropriate advice.	.000	0.71
5. The internet can help the public to know what it is like to live with a health problem.	.063	0.62
2. The internet is a reliable resource to help me understand what a doctor tells me.	-.002	0.61

Extraction Method: Principal Component Analysis.

Table 7-15 Stage Four Final items grouped by sub-scale (eHIQ-Part 1)

Sub-scale	Items	Item to total correlation	Cronbach's alpha
1.1 Attitudes towards online health information *(5 items)*	2. The internet is a reliable resource to help me understand what a doctor tells me.	0.39	0.77
	7. I would use the internet if I needed help to make a decision about my health (for example, whether I should see a doctor, take medication or seek other types of treatment).	0.64	
	6. The internet can be useful to help people decide if their symptoms are important enough to go to see a doctor.	0.60	
	8. I would use the internet to check that the doctor is giving me appropriate advice.	0.58	
	5. The internet can help the public to know what it is like to live with a health problem.	0.52	
1.2 Attitudes towards sharing health experiences online *(6 items)*	13. The internet is a good way of finding other people who are facing health-related decisions I may also face.	0.72	0.89
	9. The internet is a good way of finding other people who are experiencing similar health problems.	0.68	
	10. It can be helpful to see other people's health-related experiences on the internet.	0.72	
	14. Looking at health websites reassures me that I am not alone with my health concerns.	0.67	
	11. The internet is useful if you don't want to tell people around you (for example, your family or people at work) how you feel.	0.56	
	12. It can be reassuring to know that I can access health-related websites at any time of the day or night.	0.56	

7.5.6 Finalising the sub-scales (eHIQ-Part 2)

Using the Stage Four dataset, both the internal consistency for the sub-scales identified in Stage Three and the distributions across item response options were examined. All sub-scales with the exception of the fourth sub-scale containing items relating to the tone of a website (Cronbach's alpha= 0.68) had a value of greater than 0.74. Reliability analysis indicated the value for the fourth sub-scale would improve if Item 8 (The website left me with a lot of unanswered questions) was removed. As was the case with the eHIQ-Part 1, all eHIQ-Part 2 items were entered into an exploratory factor analysis for further analysis.

7.5.6.1 Factor analysis (eHIQ-Part 2)

The 34 items were assessed for their suitability to factor analysis. The KMO value was 0.93, exceeding the recommended value of 0.6 and the Bartlett's Test of Sphericity reached statistical significance ($p<0.01$), supporting the use of factor analysis.

Six factors with Eigenvalues above one were initially extracted. These factors explained 66.34% of the variance. Closer inspection of the Structure and Pattern matrices indicated that one factor consisted of two items (Item 8 and Item 32) which together had a Cronbach's alpha of 0.50. While Item 32 (I found the images on the website distressing) loaded on other factors, Item 8 did not appear to load on any other factor. Item 8 was therefore removed.

Following the removal of Item 8, all remaining items were entered into a further factor analysis. Five factors were extracted explaining 64.14% of the variance. The five factors were subjected to a Direct Oblimin rotation and the Structure and Pattern matrices output were used to interpret the factor structure. The five factors extracted were assessed for internal consistency and had acceptable Cronbach's alpha values of above 0.7.

Reflections on factors 4 and 5 concluded in a consensus between LK, CJ and SZ that, grouped together, items did not constitute conceptually meaningful factors. This was supported by the Scree plot which suggested only the first three factors should be extracted. As a result, the seven items in factors 4 and 5 were removed from the questionnaire (see Box 7.2 for excluded items listed). The remaining items were entered into a further exploratory factor analysis, resulting in three factors which explained 61.68% of the variance. The Pattern matrix (see Table 7-16) and the Structure matrix (see Table 7-17) were used to interpret the final factor structure.

Final sub-scales are grouped around: 2.1) *Confidence and identification*, 2.2) *Information and presentation*, and 2.3) *Understanding and motivation*. All scales had good internal reliability and are listed in Table 7-18.

7.2: Stage Four seven items removed (eHIQ-Part 2)

Items removed:

5. The website seems patronising.

11. The website provides more useful information than my doctor is likely to give me.

13. The information on the website does not apply to me.

15. It would have been useful for me to have seen the website earlier.

19. On the whole, I found the website depressing.

22. I would not want someone I care about to see the website.

32. The website is attractive

Table 7-16 Stage Four Pattern matrix (eHIQ-Part 2)

Item	Factor		
	1	2	3
19. I feel I have a sense of solidarity with other people using the website.	0.89	0.03	0.06
21. I can identify with other people using the website.	0.88	0.08	0.14
26. I feel I have a lot in common with other people using the website.	0.85	-0.01	0.08
27. The website gives me the confidence to explain my health concerns to others.	0.64	-0.10	-0.31
15. The people who have contributed to the website understand what is important to me.	0.59	0.11	-0.13
30. The website makes me more confident to discuss my health with the people around me (for example, my family or people at work).	0.52	-0.02	-0.41
25. The website gives me confidence that I am able to manage my health.	0.50	0.27	-0.26
24. I value the advice given on the website.	0.47	0.37	-0.24
13.The website prepares me for what might happen to my health.	0.44	0.04	-0.33
11. I can easily understand the information on the website.	-0.03	0.79	0.01
32. I found the images on the website distressing.	-0.17	0.77	0.09
34. The website is easy to use.	0.15	0.77	0.02
3R. The information on the website left me feeling confused.	0.02	0.74	0.01
7. The language on the website made it easy to understand.	0.02	0.72	-0.18
6. The website provides a wide range of information.	0.21	0.59	-0.13
17. I trust the information on the website.	0.11	0.57	-0.27
31. Photographs and other images were used appropriately on the website.	0.34	0.56	-0.07
1. The website encourages me to take actions that could be beneficial to my health.	-0.08	-0.01	-0.84
2. The website has a positive outlook.	-0.23	0.15	-0.72
4. The website includes useful tips on how to make life better.	-0.03	0.19	-0.71
9. I feel more inclined to look after myself after visiting the website.	0.23	-0.18	-0.69
29. The website encourages me to play a more active role in my healthcare.	0.22	0.04	-0.63
10. I have learnt something new from the website.	0.21	0.12	-0.52
28. The website helps me to have a better understanding of my personal health.	0.41	0.06	-0.49
18. I would consult the website if I had to make a decision about my health.	0.22	0.25	-0.43
22. On the whole, I find the website reassuring.	0.35	0.30	-0.36

Table 7-17 Stage Four Structure matrix (eHIQ-Part 2)

Item	Factor		
	1	2	3
19. I feel I have a sense of solidarity with other people using the website.	0.87	0.30	-0.42
21. I can identify with other people using the website.	0.84	0.31	-0.36
26. I feel I have a lot in common with other people using the website.	0.81	0.24	-0.37
27. The website gives me the confidence to explain my health concerns to others.	0.77	0.26	-0.60
30. The website makes me more confident to discuss my health with the people around me (for example, my family or people at work).	0.73	0.35	-0.67
25. The website gives me confidence that I am able to manage my health.	0.73	0.57	-0.65
24. I value the advice given on the website.	0.72	0.64	-0.66
15. The people who have contributed to the website understand what is important to me.	0.70	0.37	-0.49
13.The website prepares me for what might happen to my health.	0.62	0.34	-0.57
34. The website is easy to use.	0.40	0.81	-0.43
7. The language on the website made it easy to understand.	0.36	0.81	-0.53
11. I can easily understand the information on the website.	0.23	0.78	-0.35
3R. The information on the website left me feeling confused.	0.26	0.74	-0.36
17. I trust the information on the website.	0.45	0.74	-0.60
6. The website provides a wide range of information.	0.48	0.72	-0.52
31. Photographs and other images were used appropriately on the website.	0.56	0.70	-0.51
32. I found the images on the website distressing.	0.04	0.67	-0.18
1. The website encourages me to take actions that could be beneficial to my health.	0.36	0.36	-0.80
4. The website includes useful tips on how to make life better.	0.40	0.52	-0.79
29. The website encourages me to play a more active role in my healthcare.	0.56	0.41	-0.76
28. The website helps me to have a better understanding of my personal health.	0.69	0.44	-0.74
9. I feel more inclined to look after myself after visiting the website.	0.53	0.23	-0.73
10. I have learnt something new from the website.	0.52	0.44	-0.69
22. On the whole, I find the website reassuring.	0.63	0.58	-0.68
2. The website has a positive outlook.	0.20	0.42	-0.68
18. I would consult the website if I had to make a decision about my health.	0.52	0.53	-0.66

Table 7-18 Stage Four Final items grouped by sub-scales (eHIQ-Part 2)

Sub-scale	Items	Item to total correlation	Cronbach's alpha
2.1 Confidence and identification *(9 items)*	19. I feel I have a sense of solidarity with other people using the website.	0.79	0.92
	21. I can identify with other people using the website.	0.72	
	26. I feel I have a lot in common with other people using the website.	0.70	
	27. The website gives me the confidence to explain my health concerns to others.	0.72	
	30. The website makes me more confident to discuss my health with the people around me (for example, my family or people at work).	0.73	
	25. The website gives me confidence that I am able to manage my health.	0.75	
	24. I value the advice given on the website.	0.74	
	15. The people who have contributed to the website understand what is important to me.	0.65	
	13. The website prepares me for what might happen to my health.	0.61	
2.2 Information and presentation *(8 items)*	7. The language on the website made it easy to understand.	0.75	0.89
	11. I can easily understand the information on the website.	0.66	
	34. The website is easy to use.	0.74	
	3. The information on the website left me feeling confused.	0.65	
	6. The website provides a wide range of information.	0.68	
	17. I trust the information on the website.	0.69	
	32. I found the images on the website distressing.	0.51	
	31. Photographs and other images were used appropriately on the website.	0.67	
2.3 Understanding and motivation *(9 items)*	4. The website includes useful tips on how to make life better.	0.71	0.90
	1. The website encourages me to take actions that could be beneficial to my health.	0.67	
	2. The website has a positive outlook.	0.55	
	9. I feel more inclined to look after myself after visiting the website.	0.65	
	29. The website encourages me to play a more active role in my healthcare.	0.70	
	28. The website helps me to have a better understanding of my personal health.	0.74	
	22. On the whole, I find the website reassuring.	0.69	
	10. I have learnt something new from the website.	0.64	
	18. I would consult the website if I had to make a decision about my health.	0.64	

7.5.7 Determining sub-scale scores

A scoring system for each questionnaire part was devised using a simple summation score. Raw scale scores were calculated by summing the final response values for all the items in a given sub-scale minus the minimum raw score and dividing the value by the range between the maximum and minimum raw score for the given sub-scale. The raw scale score was then used to transform the score to a 0-100 metric by multiplying the raw scale score by 100 (See Formula 7.1).

Each scale was transformed to a 0-100 metric, where 0=low perceived value of the internet (or website) for health, and 100= high perceived benefit of using the internet (or website) in relation to health.

Formula 7.1 eHIQ sub-scale scoring algorithm (CTT derived scoring)

Formula for scoring each sub-scale =

$$\frac{Sum\ of\ scores\ of\ each\ item\ in\ scale - minimum\ raw\ score}{Maximum\ raw\ score - minimum\ raw\ score} \times 100$$

Table 7-19 and Table 7-20 provide the item numbers which belong to each sub-scale and the minimum and maximum raw scale scores.

Table 7-19 Items per sub-scale and maximum scores (eHIQ-Part 1)

Sub-scale	Sum of final response values for item numbers*	Minimum raw score	Maximum raw score
1.1 Attitudes towards online health information	2+6+7+8+12	5	25
1.2 Attitudes towards sharing health experiences online	9+10+11+12+13+14	6	30

Table 7-20 Items per sub-scale and maximum scores (eHIQ-Part 2)

Sub-scale	Sum of final response values for item numbers*	Minimum raw score	Maximum raw score
2.1 Confidence and identification	13+15+19+21+24+25+26+27+30	9	45
2.2 Information and presentation	3+6+7+11+17+31+32+34	8	40
2.3 Understanding and motivation	1+2+4+9+10+18+22+28+29	9	45

* Item numbers reflect item numbers used in Stage Four administration. For final questionnaire item numbers and respective scoring tables see Appendix 7.25.

7.5.7.1 Stage Four eHIQ score distributions

Descriptive statistics for the final eHIQ sub-scales are shown in Table 7-21. The eHIQ sub-scale scores follow a relatively normal distribution across the study sample. This is visually supported through histograms of sub-scale scores (see Figures 1-4, Appendix 7.26). Results do however show the 2.2) *Information and presentation* sub-scale to have a slightly skewed pattern towards the higher end of the scale (positive feelings towards information and presentation). The sub-scale is also kurtotic. In contrast to the skewness (i.e. how symmetrical the distribution is around the mean), kurtosis gives an indication of how acute the peak is around the mean. These observations indicate a non-normal distribution of scores. See Figure 7-1 for visual evidence of non-normality in distribution for this scale. This pattern was somewhat expected as the websites which participants were asked to view were of good-high quality. The reference measures also show some non-normality of distribution among the sample (see Figures 5 and 6, Appendix 7.26). Therefore statistical analyses for scale scores in this study are reported using both parametric and non-parametric statistics. See specific analyses for test details.

Scores were examined for floor and ceiling effects. High frequencies at the limits of the scales can suggest a lack of sensitivity at the extreme points. While it is preferable to cover a wide score range, large percentages at the extreme ends of the scale is undesirable. No scales exhibited floor or ceiling effects, with the sub-scale 1.2) *Attitudes towards sharing health experiences online* marginally exceeding the ceiling 5% cut off point with 5.9% of respondents achieving a score of 100 (see Table 7-22).

Table 7-21 Stage Four scale score descriptive characteristics

Sub-scale	N	Minimum	Maximum	Mean	SD	Skewness	Kurtosis
1.1 Attitudes towards online health information	169	0.0	100.0	63.31	16.90	-0.775	0.850
1.2 Attitudes towards sharing health experiences online	169	0.00	100	70.69	16.30	-0.68	1.69
2.1 Confidence and identification	165	0.0	100.0	58.17	17.54	-0.739	0.964
2.2 Information and presentation	167	0.0	100.0	73.18	13.93	-1.69	6.75
2.3 Understanding and motivation	167	8.33	97.22	61.23	16.07	-0.889	1.75
HINTS	163	1	4	2.13	0.53	0.66	1.81
Access to Quality Information (8 items)*	168	1	4.88	3.72	0.57	-1.51	4.61
Access to Quality Information (6 items)	169	1	5	3.61	0.61	-1.27	3.29

*Mean scale score (Web Trust Questionnaire - 8 items, 5 point response scale)

Figure 7-1 Scale 2.2: 'Information and presentation' scale score distribution

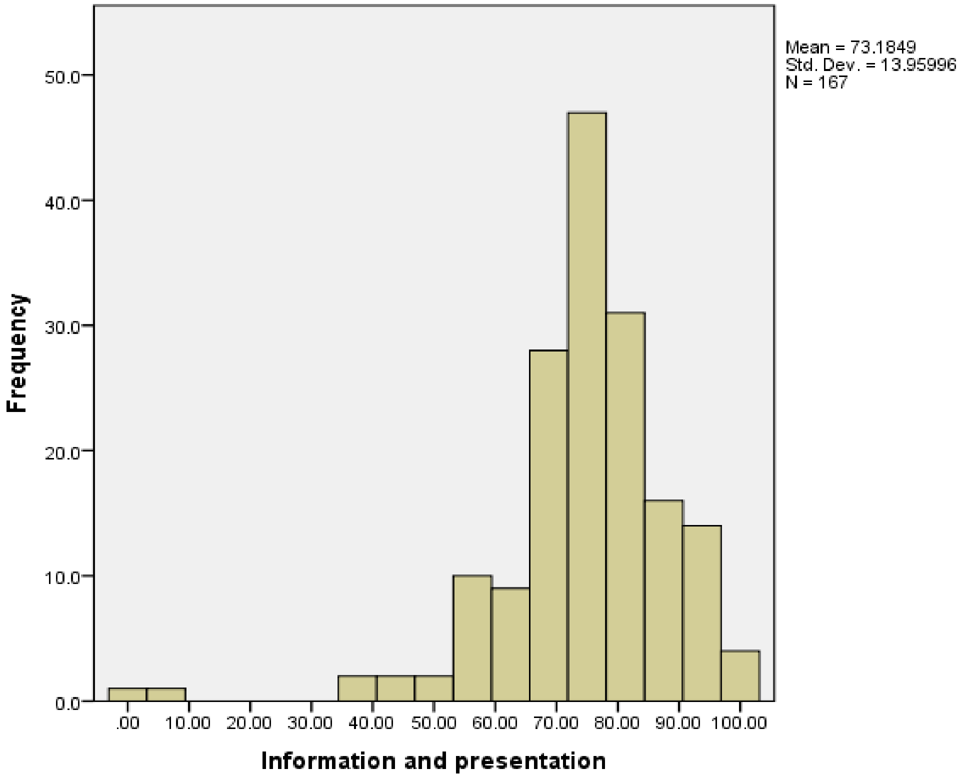


Table 7-22 Stage Four floor and ceiling effects

Sub-scale	Floor (Score=0)		Ceiling (Score=100)	
	n	%	n	%
1.1 Attitudes towards online health information	1	0.6	1	0.6
1.2 Attitudes towards sharing health experiences online	1	0.6	10	5.9
2.1 Confidence and identification	1	0.6	1	0.6
2.2 Information and presentation	1	0.6	1	0.6
2.3 Understanding and motivation	0	0	0	0

Floor and ceiling >5% highlighted

7.5.8 Validation

In this section the relationships between the sub-scales and a range of population characteristics are examined to identify potential covariate factors related to population characteristics. Relationships between the eHIQ and the selected reference measures are also examined to assess convergent validity.

7.5.8.1 Population factors

7.5.8.2 Gender

Parametric (t-test) and non-parametric tests (Mann-Whitney U test of Significance) were carried out to investigate if there were significant differences according to gender and eHIQ scores. No significant differences were found between sex and all sub-scale scores (see Table 7-23). Descriptive statistics are displayed in Appendix 7.27.

7.5.8.3 Age

Parametric (Pearson correlation) and non-parametric (Spearman Correlation) tests were carried out to investigate if there were significant differences in eHIQ score according to age. No significant differences were found between age and sub-scale scores in all parametric tests. Non-parametric tests indicated there was a significant positive relationship between scale 2.1) *Confidence and identification* and age however the relationship was extremely weak (Spearman correlation=0.18, p=0.038) (see Table 7-24).

7.5.8.4 Recruitment phase

Parametric tests (One way ANOVAs) were carried out to test for significant differences among the eHIQ scales according to the method of recruitment. The significance for the

Homogeneity of variance statistic F was greater than 0.05 in all calculations indicating it was valid to conduct ANOVAs.

There were significant differences on the sub-scale 1.2) *Attitudes towards sharing health experiences online* between the three recruitment phases [F (2,166) = 4.60, p=0.011]. Post-hoc comparisons using Tukeys indicated that there were significant differences between those in Phase 1 (Mean=65.51, SD=14.89) and Phase 3 (Mean=76.85, SD=15.17). Eta^2 was equal to 0.05 (5%) indicating a small effect. Non-parametric tests (Kruskal-Wallis ‘k independent samples) supported this finding with significant differences on the eHIQ sub-scale 1.2) *Attitudes towards sharing health experiences online* between the three recruitment phases (p=0.007) (see Table 7-25). This may have been due to the sample size in Phase 3. Descriptive statistics are displayed in Appendix 7.28.

Table 7-23 Relationship between sex and eHIQ sub-scales

Sub-scale	**Sex**	**n**	**Mean (SD)**	**Parametric**		**Non-parametric***
				t	**p**	**p**
1.1 Attitudes towards online health information	Male	59	61.0 (17.7)	-1.781	0.077	0.172
	Female	106	65.7 (15.3)			
1.2 Attitudes towards sharing health experiences online	Male	59	69.8 (17.3)	-0.653	0.514	0.812
	Female	106	71.5 (15.7)			
2.1 Confidence and identification	Male	58	60.3 (17.1)	1.214	0.227	0.271
	Female	103	57.2 (17.3)			
2.2 Information and presentation	Male	58	73.8 (13.9)	0.212	0.833	0.756
	Female	105	73.3 (13.9)			
2.3 Understanding and motivation	Male	58	63.2 (14.4)	1.011	0.313	0.441
	Female	105	60.6 (16.5)			

*Non-parametric test: Mann-Whitney U test of Significance

Table 7-24 Relationship between age and eHIQ sub-scales

Sub-scale	**n**	**Parametric**		**Non-parametric**	
		Pearson r	**p**	**Spearman Correlation**	**p**
1.1 Attitudes towards online health information	142	-0.097	0.251	-0.093	0.269
1.2 Attitudes towards sharing health experiences online	142	-0.125	0.139	-0.159	0.059
2.1 Confidence and identification	141	0.117	0.167	0.175	0.038*
2.2 Information and presentation	142	-0.066	0.436	-0.063	0.459
2.3 Understanding and motivation	142	-0.040	0.633	-0.002	0.979

*Significant at <0.05 level (2 tailed)

Table 7-25 Relationship between recruitment phase and eHIQ scales

Sub-scale	Phase	n	Mean (SD)	Parametric		Non-parametric
				F	P	P
1.1 Attitudes towards online health information	1	47	61.17 (16.29)	2.03	0.135	0.136
	2	95	62.74 (17.35)			
	3	27	69.07 (15.51)			
1.2 Attitudes towards sharing health experiences online	1	47	65.51 (14.89)	4.60	0.011*	0.007*
	2	95	71.49 (16.67)			
	3	27	76.85 (15.17)			
2.1 Confidence and identification	1	47	59.22 (15.53)	1.41	0.249	0.145
	2	94	56.44 (18.01)			
	3	24	62.85 (19.01)			
2.2 Information and presentation	1	47	74.47 (13.02)	0.31	0.732	0.689
	2	95	72.50 (13.72)			
	3	25	73.38 (16.73)			
2.3 Understanding and motivation	1	47	62.29 (15.47)	0.16	0.856	0.922
	2	95	60.94 (15.84)			
	3	25	60.33 (18.45)			

*Significant at <0.05 level (2 tailed)

7.5.8.5 Convergent validity

Pearson's correlation coefficients were used to test the relationships between the eHIQ sub-scale scores and the selected reference measures (see Table 7-26). Correlations between the eHIQ-Part 1 sub-scales and the HINTS single item (In general, how much would you trust information about health or medical topics on the internet?) were weak to moderate ranging from -0.35 to -0.52. Correlations between the eHIQ-Part 2 sub-scales and the 'Access to quality information' items were moderate to good ranging from 0.74 to 0.84 (eight item scale) and 0.75 to 0.76 when excluding the two overlapping items which are contained in both measures. Results confirm expectations that the measures are significantly related under the common construct but are sufficiently divergent in terms of specific content.

Table 7-26 Relationship between eHIQ sub-scales and reference measures

Sub-scale	**Relationship with (Pearson r)**			
	n	**HINTS**	**Quality information (8 items)**	**Quality information (6 items)**
1.1 Attitudes towards online health information	163	-0.52*	-	-
1.2 Attitudes towards sharing health experiences online	163	-0.35*	-	-
2.1 Confidence and identification	165	-	0.74*	0.75*
2.2 Information and presentation	167	-	0.84*	0.75*
2.3 Understanding and motivation	167	-	0.77*	0.76*

*Significant at <0.05 level

7.5.8.6 Reliability

To establish reliability of the measure the test-retest procedure was used. Participants were invited to complete the eHIQ on two separate occasions with a two week interval. Participants were asked to browse the same website at both time points.

Of the 170 sample, 143 were asked to complete a second questionnaire (those in the Phase 3 group were not asked to complete a second questionnaire). Of the 143 people asked, 90 (62.94%) people responded. Due to a computer malfunction which meant it was impossible to link up the two surveys to the same respondent, 22 responses from the MND participants had to be disregarded. Seven of the 90 respondents did not complete the eHIQ-Part 2. The level of agreement tested ranged from 0.76 to 0.91 indicating good test-retest reliability for all sub-scales (see Table 7-27).

Table 7-27 Intra-class correlation coefficient for final eHIQ sub-scales

eHIQ sub-scale	**n**	**ICC***
1.1 Attitudes towards online health information	68	0.85
1.2 Attitudes towards sharing health experiences online	68	0.76
2.1 Confidence and identification	61	0.89
2.2 Information and presentation	61	0.79
2.3 Understanding and motivation	61	0.91

* Absolute agreement

7.6 Discussion

Stage Four primarily aimed to finalise the eHIQ sub-scales and test the psychometric properties of the instrument. The analyses outlined in this chapter indicated that the eHIQ-Part 1 has two sub-scales: 1.1) Attitudes towards online health information and 1.2) Attitudes towards sharing health experiences online, while the eHIQ-Part 2 has three sub-scales: 2.1) Confidence and identification, 2.2) Information and presentation and, 2.3) Understanding and motivation.

7.6.1 Translatability

Before Stage Four data collection, a translatability assessment was carried out. The translatability assessment was successful in highlighting a number of potential ambiguities and cultural or linguistic problems within the eHIQ. Minor adjustments were made to instructions and items which may ease future translations. It is important to note that all recommendations from translation experts were considered, however, they were not all implemented in order to balance the risk of improving translatability at the cost of excluding important concepts in the source language.

7.6.2 Psychometric properties

The final eHIQ sub-scales are valid and reliable displaying good internal consistency and moderate to good correlations with relevant reference measures. Scores on the eHIQ-Part 1 were compared with the HINTS single item and eHIQ-Part 2 sub-scales were compared to scores on the *Access to Quality Information* scale from the Web Trust questionnaire (Harris et al., 2011). It was hypothesised that the respective scores would be related to the eHIQ sub-scales providing they were measuring similar constructs. If this were the case, scores were predicted to have moderate correlations with the eHIQ sub-scales. Strong correlations can indicate scales are very similar and the new eHIQ sub-scales were unnecessary. Results indicated weak-moderate correlations with the eHIQ-Part 1 and moderate correlations for the eHIQ-Part 2. This was considered acceptable evidence of convergent validity.

Various population characteristics were investigated to see if items performed differently across groups. The two part measure did not perform differently between genders, however some evidence exists to suggest that one sub-scale (2.1 *Confidence and identification*) behaved slightly differently for age. This positive relationship was very weak and is therefore unlikely to have any meaningful implications.

Some significant differences were found between scores for scale 1.2) *Attitudes towards sharing health experiences online* among those who were recruited via mail and those via the internet. This may be linked to the differences found between these groups for the frequency of internet use. Scores for scale 1.2 for those recruited in Phase 1 (mail recruitment) were less positive than those in Phase 3 (online recruitment). This may be due to the sample size in Phase 3, however it is possible that those who were recruited online may be more positive to online experiential information. This assertion would need further testing with a larger sample size. Overall results demonstrated that the sub-scales largely performed in a similar manner among the identified demographic variables. This supports the use of the eHIQ across these characteristics.

Test-retest reliability achieved a good response rate of 58.4%. Unfortunately this sample size became smaller due to an IT malfunction meaning the linking identifiers were not available. Among those respondents who could be linked to their initial questionnaire there appeared to be strong test-retest reliability across all sub-scales. It is recommended that future attempts to assess test-retest reliability should ask participants to input some identifier (for example, their email address) on the Time 2 survey in case of technical difficulties.

Further difficulties with examining test-retest reliability were encountered. Aside from the technical malfunction discussed, this particular questionnaires' retest was susceptible to external influences. Participants were asked to explore a specific health-related website in order answer the eHIQ-Part 2. Although care was taken to ensure the external website and its content had not recognisably changed during the two time points, it was impossible to guarantee the websites content had not changed. Furthermore, many websites take a considerable time to explore in their entirety. It was therefore possible that the respondents changed their opinion of aspects of the website having engaged with a different section of the same website on the second occasion. This was reported by one participant, however, it is difficult to speculate if it occurred more than once. If participants were restricted to one viewing only (i.e. on the first occasion), it would have been unrealistic to expect them to complete the questionnaire accurately on the second occasion from recall. In future studies which may wish to test the external reliability of the eHIQ (for example, in a different population group), it is recommended that the researcher uses a static website which is under their control. One further obstacle to retesting the eHIQ-Part 2 concerned items which were unlikely to elicit the same response on the second occasion (for example, ‘I have learnt something new from the website.’) Despite these concerns however, test-retest values were found to be desirable.

7.6.3 Recruitment and sample

In comparison to other survey modes, web surveys are economical as costs remain the same regardless of sample size (Sue and Ritter, 2012). Data collection may also be carried out at a faster pace using electronic surveys compared to traditional forms of data collection (for example, pen and paper). It can be difficult however to successfully target and successfully recruit the desired population and recruitment was considered poor over all phases in Stage Four. Nonetheless, online recruitment rates should not be treated in the same manner as postal recruitment. Phase 1 and 2 explored the feasibility of two recruitment strategies for online questionnaires. Difficulties are discussed below.

Phase 1 used a postal recruitment strategy asking a pre-defined population to take part in the study. Selecting potential participants from the GP database enabled less reliance on self-report as participants with a known condition, age and gender could be targeted and an accurate response rate could be calculated. It is important to note that the accuracy of the information was dependent on the reliability of the database. For example, a number of people called to say that they did not wish to stop smoking despite being labelled as such on their GP database.

A response rate of 6.1 % was considered extremely low despite knowing the low rates of response in previous studies recruiting from primary care (NHS_England, 2013a). A number of factors may be considered as contributing to this poor result. Asking participants to complete an online survey required access to the internet and some degree of computer literacy. It was not possible to screen for these requirements from the GP database. On agreeing to participate, patients were asked to access the questionnaire via the University of Oxford's departmental website. This was an additional hurdle for participants who are typically asked to complete a hard copy questionnaire and return it in a stamped addressed envelope. Asking participants to access the University's webpage first was designed to add legitimacy to the request, however, it may have been bothersome or difficult for participants to carry out and a lower response rate may therefore have been reasonably expected. On reflection, a more feasible solution may have been to ask participants to email LK in order to obtain an electronic link to the questionnaire sent directly to their inbox. However, the merits of such a system would need to be further explored as some participants could view this as an 'additional' step to accessing the questionnaire and therefore potentially more burdensome.

Of those who were recruited through GP practices, people with asthma had a higher response rate than those wishing to stop smoking. Targeting a smoking cessation population was difficult due to the multiple Read codes associated with the population group (see Table 7-4, Section 7.3.7.2). Although considerable effort went into ensuring the optimal Read codes were selected (for example, discussions with practice nurses), it is possible that using other Read codes may have been more successful. In addition, this study was a purely methodological one which arguably had little incentive for participants to take part. For example, gaining access to a new website may have increased enthusiasm to take part when compared to a study aimed at developing a questionnaire.

Finally, the overall experience of recruiting participants via GPs for an online survey through postal research invitations was both time consuming and costly. A considerable amount of time went into approaching and recruiting GP practices to take part, collating research invitation packs and reminder letters and following up practices to ensure timely distribution. Costs included were administrative and postal expenses.

Phase 2 recruited participants by emailing research invitations to potential participants and through open research advertisements on publicly available websites. Internet users may attribute less legitimacy to internet recruitment compared to traditional postal recruitment due to experiences of junk mail (spam), scam emails and security threats from unknown sources (Koo and Skinner, 2005). Where possible, recruitment material was therefore distributed or posted by organisational administrators and webmasters to add authenticity to recruitment requests (Koo and Skinner, 2005, Morgan et al., 2013, Murray et al., 2009). In addition to adding authenticity to emailed research invitations, research materials were sent by email list administrators to decrease the risk of the invitation going to junk mail boxes. This facilitated maintaining participant confidentiality before participant consent had been given. Although many gatekeepers were extremely helpful in this study, a considerable amount of time was spent gaining access to potential organisations which could help in the research (for example, 17 MS Society branches were contacted in Phase 2 yet only four branches emailed research invitations to their members). Branch representatives who responded to the request for help with the study, but declined to take part, gave reasons such as the lack of an email membership list and that many of their members did not use the internet very frequently.

Approximate response rates were estimated for those who were emailed a recruitment pack. This estimate was reliant on the email administrator giving an accurate number of people included on their email list and did not account for emails which may have been forwarded by participants to others they thought may like to take part or email addresses on membership lists which may have been no longer in use.

Estimating the response rate for *open* recruitment to an online questionnaire can lead to spurious results depending on the method chosen (Eysenbach, 2004). Internet users browse and discard web pages at speed. Calculating the response rate solely through counting the number of people (through IP addresses or new survey entries) who may have viewed a research advertisement as the denominator is inaccurate (Dobrow et al., 2008) and all advertisement viewers will not be eligible to take part in the study. Those who were recruited using an online advertisement on a health website however had immediate access to the internet, were likely to have some computer competency and have some interest in online health information. Online recruitment however remained disappointing and reflected experiences in other studies. Previous research has also highlighted the difficulty in recruiting participants online using both open and direct recruitment methods where direct recruitment yielded a response rate of 0.24% and open recruitment yielded only 14 participants despite the offer of monetary (evoucher) rewards (Koo and Skinner, 2005).

Questionnaire completion rates can be a more informative and pragmatic estimate for online recruitment (Dobrow et al., 2008). Completion rates were calculated using the number of people who consented to take part in the study as the denominator and were reasonably good (71.1%). This suggested the questionnaire itself was not a deterrent to participation.

The general premise behind calculating response rates concerns the generalisability of the data. Online recruitment response or completion rates however are unlikely to give an indication of how representative a sample may be to the wider population. Using these figures as an indication of the generalisability of results is therefore limited and, while the internet will continue to be a good tool for data collection, representativeness of the study sample can be compromised (Eysenbach and Wyatt, 2002). This particular problem was flagged in the previous chapter as the data were collected using open recruitment methods only. Stage Four (Phase 1) aimed to achieve a representative sample for the asthma and smoking cessation group in relation to age and gender. Bias and self-selection was minimised in Phase 2 through recruiting from a wide range of populations, websites and email lists (Reips, 2002). Overall, while Stage Four explored the feasibility of recruitment strategies, it is important to be mindful that a representative population for this particular research should be confined to those who use the internet and not the general population.

With these considerations in mind, the Stage Four sample characteristics were thought to be satisfactory. More women than men were recruited in this study. This was somewhat expected as women are more likely to access health information online than men (Fox et al., 2000, Fox, 2011b). As this questionnaire was designed for use amongst people who use the internet, the gender profile of the sample is therefore likely to be a true reflection of the population of interest. Participants recruited via the internet were younger than participants recruited through the research invitation mail outs. This is also comparable to populations sampled in previous internet related studies (Graham et al., 2008, Etter and Perneger, 2001).

The sample was predominantly white and future analysis of the questionnaire would benefit from testing among a more ethnically diverse population. Future studies would need to implement an appropriate recruitment strategy to get ethnic diversity in the sample to ensure it is adequate and balanced (Im, 2005).

7.6.4 Challenges to internet survey research and design

Using the internet as a research and recruitment tool presented a number of technical and design challenges. One of the biggest challenges to this research was maintaining control in the research setting. To ensure the integrity of the database, it was important to establish that each participant only took part once in each time point (Etter and Perneger, 2001, Reips, 2002). It was possible to identify more than one entry per participant using the Qualtric's survey software (IP addresses) and, where necessary, it was possible to investigate participants details (email addresses and demographic information) where duplicate IP addresses arose. This process indicated no entries were likely to have been submitted by the same participant. The lack of monetary incentives may have discouraged multiple completion, however research has shown multiple responses are low risk in internet based experiments (Koo and Skinner, 2005, Reips, 2002).

This study also aimed to restrict recruitment to participants within the UK. Geographically restricting participation is theoretically possible through checking IP addresses in real time, this however would have been hard and expensive to carry out in this small scale study. Postings and advertisements in Phase 2 were therefore restricted to UK websites in order to reduce the potential of a more global catchment (Koo and Skinner, 2005) and participants (online recruitment) were asked to confirm they were resident in the UK while completing consent.

Administering the questionnaire and collecting data took careful planning and preparation over several stages. Previous chapters have documented the use of cognitive debrief interviews to pilot test the questionnaire and the use of BOS software to administer the questionnaire. Changing the survey software after Stage Three facilitated greater flexibility in design and improved user experience through the availability of additional features such as a progress bar. Qualitric's survey software had a license fee making administration more expensive. Nonetheless, having the right software enabled some difficulties to be overcome.

As has been found in previous research (Morgan et al., 2013), participants reported some difficulty locating the 'Continue' button at the bottom of the screen (this may have required scrolling down on some computer monitors). This was rectified by relocating the 'Continue' higher on the screen and increasing the size of the 'Continue' button.

Participant feedback in Phase 1, indicated the participant information text was too lengthy. While participant information was necessary, including long text is contrary to usability recommendations for webpages (Morgan et al., 2013). The appearance of the text was therefore amended in Phase 2 to display participant information subheadings and drop down answers when selected.

7.7 Conclusion

This chapter outlines the final eHIQ sub-scales. All sub-scales show good internal and external reliability. Convergent validity of the sub-scales are supported through moderate to good correlations with relevant reference measures. Minor adjustments to wording of the questionnaire resulted in an instrument which should be suitable for translation into the six identified languages and is likely to be amenable to translation in further languages and cultures. Various forms of offline and online recruitment were examined which should inform recruitment strategies future studies.

CHAPTER 8: The eHealth Impact Questionnaire: Summary index scores and short form questionnaire

8.1 Introduction

Instruments measuring more than one domain (or factor) of a wider construct provide a detailed assessment of the various dimensions contributing to the overall construct. For example, the eHIQ-Part 1 consists of two factors, and therefore has two respective scores. Together these scores are thought to measure attitudes towards using online health information. Similarly, the eHIQ-Part 2 consists of three factors, and therefore three respective scores. Together these scores measure the impact of using a specific health-related website. These factor scores can be used to compare each sub-scale with a variable of interest. Multiple scores however can be difficult to manage in analyses and it may be desirable to make one overall comparison of the overall construct to a specific variable. Multiple comparisons may also increase the risk of 'significant findings' due to chance and reducing the number of factors being compared in hypothesis testing can reduce chance (Ware et al., 1994, Ware et al., 1995). Consequently the possibility of creating a higher order factor has been advocated so that an overall summary index score or scores can be generated (Nunnally and Bernstein, 1994, Ware et al., 1995).

Factor analysis aims to explain the maximum amount of variance within a set of items using the smallest number of latent factors. Implicit in all oblique factor rotations are higher-order factors (Gorsuch, 1983). As outlined in previous chapters, the eHIQ items were subjected to an oblique rotation identifying two correlated factors in Part 1 and three correlated factors in Part 2. These factors are termed first order factors and a factor correlation matrix for each questionnaire part was produced. These first order factor correlations are suitable for entry into higher order factor analyses to identify second order factors often referred to in health services research as higher order factors (Nunnally and Bernstein, 1994, Kline, 2000). Creating summary index scores for health-related assessment tools has been common practice in previous studies (Morley et al., 2013, Jenkinson et al., 2013, Ware et al., 1995, Ware et al., 1994). In the case of the eHIQ, these second order factor scores can aid interpretation in relation to overall attitudes towards online information and overall impact of using a specific health-related website. Consequently a higher order factor analysis was carried out for the eHIQ-Part 1 and the eHIQ-Part 2 to see if the sub-scales in each part of the questionnaire could be aggregated to create two respective summary index scores.

In addition to creating an overall summary score for each part of the questionnaire, the feasibility of creating a short form questionnaire was explored. Short form questionnaires have been developed for a large number of widely used questionnaires and many have been found to be acceptable and informative providing similar results to their original form (De Bruin et al., 1994, Ware and Sherbourne, 1992, Bohlmeijer et al., 2011). The eHIQ-Parts 1 and 2 contain 37 items in total. In some studies where other measures are administered in conjunction with the eHIQ, completing 37 items may be considered burdensome for the participant. To ease responder burden and reduce the possibility of missing data, a shorter version of the eHIQ was sought.

8.1.1 Objective

These analyses aimed to identify a higher order factor solution for the eHIQ Parts 1 and 2 to enable the calculation of summary index scores. Analyses also sought to explore the feasibility of creating short form questionnaires for each part.

8.2 Methodology

8.2.1 Higher order factor analysis

Data from the first order factor analysis carried out in Chapter 7 was used in these analyses. The first order factor by factor correlation matrix for the eHIQ Part 1 (see Table 8-1) and the eHIQ Part 2 matrix (see Table 8-2) provided input for their respective higher order factor analysis (Gorsuch, 1983).

Similar to the item correlation matrices (derived from raw data) which were used in the first order factor analyses, the factor correlation matrices were inputted into SPSS using the syntax editor. Use of the factor correlation matrices was preferred over the use of first order factor scores as variance which did not contribute to the first order factors may still have been useful in explaining the variance within a second order factor (Gorsuch, 1983).

Although tests for significance on the correlation matrix were used in the first order factor analysis (Bartlett’s Test of Sphericity testing for relationships among the items and the KMO statistic for sampling adequacy), it was unsuitable to use such tests in higher order factor analyses as correlation coefficients vary according to the rotation used in the primary analysis and the sample size (Gorsuch, 1983, Evans, 1999).

Table 8-1 Factor correlation matrix (eHIQ- Part 1)

Factor	Scale 1.1	Scale 1.2
Scale 1.1	1.000	0.524
Scale 1.2	0.524	1.000

Table 8-2 Factor correlation matrix (eHIQ- Part 2)

Factor	Scale 2.1	Scale 2.2	Scale 2.3
Scale 2.1	1.000	0.335	-0.523
Scale 2.2	0.335	1.000	-0.476
Scale 2.3	-0.523	-0.476	1.000

In case more than one factor was identified, decision rules for retaining factors were similar to those used in previous primary factor analyses and included: having an eigenvalue of greater than one and examining Catell's scree plot to identify a sharp break in the descending curve of eigenvalues (Gorsuch, 1983). The reliability of each higher order factor was assessed using the Cronbach's alpha statistic where a value of > 0.7 was considered satisfactory.

Summary index scores were calculated through summing the sub-scale scores and dividing by the number of sub-scales to get a 0-100 metric.

8.2.2 The short form questionnaire

High levels of internal consistency can indicate item redundancy and a smaller number of items may be chosen to represent the construct in question. To explore the possibility of creating a short form questionnaire, the two most highly correlated (Pearson's r) items to each sub-scale were selected. Due to the relatively low number of sub-scales in each part of the questionnaire, two items for each sub-scale were chosen to maximise accuracy and maintain face validity.

The selected items were collated and their internal consistency was examined (Cronbach's alpha, >0.7). Items were also entered into a factor analyses to ensure the appropriateness of combining the items (i.e. to check the items would group as one factor). The level of agreement between each summary score and their respective short form score was tested using the intra-class correlation coefficient (ICC). Coefficients above 0.70 were considered psychometrically satisfactory and values of above 0.80 deemed particularly good (Nunnally and Bernstein, 1994).

8.3 Results

8.3.1 Summary index score (eHIQ- Part 1)

The factor correlation matrix for the eHIQ Part 1 was entered into a factor analysis. One factor with an eigenvalue of value greater than one was extracted, explaining 76.2% of the variance. No rotation was needed as all sub-scales loaded sufficiently on the factor extracted (see Table 8-3, eHIQ-Part 1).

The higher order factor, referred to here as *Attitudes towards the use of the internet in relation to health* was assessed for internal consistency achieving a Cronbach's alpha of 0.78 (n=169). This supported summing all eHIQ-Part 1 sub-scales to create a total index score. After the summed score for the higher order factor had been standardised to 0-100 metric, the mean summary score was 67.0 (SD 15.04, 95% CI, 64.72-69.28, n=169) (see Table 8-4, eHIQ-Part 1).

8.3.2 Summary index score (eHIQ- Part 2)

The factor correlation matrix for the eHIQ Part 2 was subsequently entered into a factor analysis. One factor with an eigenvalue of value greater than one was extracted, explaining 63.1% of the variance. No rotation was needed as all sub-scales loaded sufficiently on the factor extracted (see Table 8-3, eHIQ-Part 2).

The higher order factor, referred to as *Overall impact of using a specific health-related website* was assessed for internal consistency achieving a Cronbach's alpha of 0.87 (n=163). The mean summary score 64.22 (SD 14.20, 95% CI, 62.02-66.41, n=163) (see Table 8-4, eHIQ-Part 2).

The overall summary index scores for both questionnaire parts had a weak to moderate correlation (eHIQ-Part 1 and eHIQ-Part 2, r=0.33, n=163, p<0.001).

Table 8-3 Higher order factor loadings

First order factor	Factor loading on higher order factor
eHIQ-Part 1	
Scale 1.2	0.873
Scale 1.1	0.873
eHIQ-Part 2	
Scale 2.3	-0.854
Scale 2.1	0.779
Scale 2.2	0.747

Table 8-4 eHIQ Summary scores and short form scores

Sub-scale	N	Minimum	Maximum	Mean	SD	Skewness	Kurtosis
eHIQ-Part 1	169	0.00	97.92	67.0	15.04	-0.79	1.93
eHIQ-Part 2	163	2.78	99.07	64.22	14.20	-1.30	3.65
SF eHIQ-Part 1	169	0.0	100.0	63.54	17.96	-0.35	0.22
SF eHIQ-Part 2	167	0.0	100.0	65.19	15.29	-1.44	4.01

8.3.3 The Short Form questionnaire

The two most highly correlated items to each respective sub-scale were selected for the eHIQ-Part 1 to create a short form version of the questionnaire (SF eHIQ-Part 1) (see Table 8-5). The reduced short form questionnaire demonstrated good internal consistency (Cronbach's alpha= 0.74, n=169) and, when entered into a factor analysis, all items loaded on one factor explaining 56.90% of the variance. The mean short form score for the eHIQ-Part 1 was 63.54 (SD 17.96, CI 60.81-66.26, n=169).

Table 8-5 Short Form items and correlations with respective sub-scales (eHIQ-Part 1)

Item	**r***
7. I would use the internet if I needed help to make a decision about my health (for example, whether I should see a doctor, take medication or seek other types of treatment).	0.812
8. I would use the internet to check that the doctor is giving me appropriate advice.	0.766
10. It can be helpful to see other people's health-related experiences on the internet.	0.813
13. The internet is a good way of finding other people who are facing health-related decisions I may also face.	0.817

*All significant at 0.01 level

Similarly the two most highly correlation items to their respective sub-scale in the eHIQ-Part 2 were selected (SF eHIQ-Part 2) (see Table 8-6). The reduced questionnaire also had good internal consistency (Cronbach's alpha= 0.86, n=167) and, when entered into a factor analysis, all items loaded on one factor explaining 57.88% of the variance. The mean short form questionnaire score for the eHIQ-Part 2 was 65.19 (SD 15.29, CI 62.86-67.53, n=167).

Table 8-6 Short Form items and correlations with respective sub-scales (eHIQ-Part 2)

Item	**r***
19. I feel I have a sense of solidarity with other people using the website.	0.843
25. The website gives me confidence that I am able to manage my health.	0.805
7. The language on the website made it easy to understand.	0.812
34. The website is easy to use.	0.803
28. The website helps me to have a better understanding of my personal health.	0.805
4. The website includes useful tips on how to make life better.	0.777

*All significant at 0.01 level

Both the short form scores had very high Intra Class Correlation Coefficient (ICC) scores with their respective overall summary scores (SF eHIQ-Part 1 and eHIQ-Part 1 summary score (ICC=0.91, n=169, p<0.001) and SF eHIQ-Part 2 and eHIQ-Part 2 summary score (r=0.96, n=163, p<0.001)).

8.4 Discussion

The aim of this chapter was to identify a higher order factor solution for the eHIQ-Parts 1 and 2 and explore the feasibility of creating two short form questionnaire parts. Higher order factors provide a broader scope of the overall construct being measured allowing increased generalisability, while the initial first order factors provide a narrow scope, but more accuracy (Gorsuch, 1983). Therefore, identifying a higher order factor solution was sought to support summing the sub-scales to obtain a single summary index score for each questionnaire part. A short form questionnaire was sought to reduce responder burden and decrease the chance of missing data.

The results outlined in this chapter provide evidence that it is appropriate to create a summary index score for the eHIQ-Part 1 and the eHIQ-Part 2. Higher order factor analyses confirmed the sub-scales within each questionnaire part load together as single factors. The reliability, indicated by high Cronbach's alpha values (≥ 0.87), of the summary scales supported the internal consistency of the scales. The creation of summary scores will potentially aid interpretation and analyses in future studies as comparisons of the overall constructs with other relevant variables can be easily made. As predicted, attitudes towards the use of the internet in relation to health were significantly correlated to overall impact of using the specific health-related website viewed. Their correlation however was relatively weak.

Analyses also found an acceptable short form questionnaire for Parts 1 and 2. Internal reliability for both short form scales was confirmed using Cronbach's alpha values (≥0.74).

Validity of the short form questionnaires were supported through the high, significant correlations between the summary score for each questionnaire part and their respective short form questionnaire (r ≥0.93). Face validity is supported as items within the short form questionnaires also appear highly relevant to the two overarching constructs. The short form questionnaires provide an alternative method of assessment where it is thought the full questionnaire is too long or where it may be an appropriate secondary measure. As is the case with other established short form questionnaires (De Bruin et al., 1994, Ware and Sherbourne, 1992, Bohlmeijer et al., 2011), the use of the short form eHIQ appears to be informative as it provides similar results to its parent form. Therefore, while using the full length version of the eHIQ questionnaire can potentially provide slightly more accuracy in measurement, the short form questionnaire is recommended for use where responder burden may be a concern in future studies. Decisions on whether to use the short form questionnaire over the long form questionnaire however should be based upon the purpose and priorities of the research study as use of the short form questionnaire will limit analysis (i.e. analysis cannot be carried out at the sub-scale level).

In summary, using either the overall summary index scores or the short form questionnaire will decrease the role of chance where multiple comparisons are necessary in future research. Use of the short form questionnaire will also decrease responder burden, however, its suitability of use should be assessed according to the purpose of the research study in question.

CHAPTER 9: Creating interval level scaling for the eHealth Impact Questionnaire

9.1 Introduction

Classical test theory (CTT) methods have thus far been adopted to establish the internal structure of the eHIQ-Parts 1 and 2. Instruments developed using CTT produce scores which are theoretically comprised of two variables: the true score and the error score (Nunnally and Bernstein, 1994). The error score is random and, in CTT, it is postulated that it has a mean of zero across all items and is uncorrelated to the true score (Nunnally and Bernstein, 1994, Hambleton and Jones, 1993). These assumptions enable the observed response score across items to be summed to create a total score indicating more or less of the construct in question (i.e. For eHIQ-Part 1: more or less positive views of online health information; and for eHIQ-Part 2: more or less positive impacts of using a specific health-related website). From a theoretical perspective, this simple summation implies that items within each scale are of equal difficulty to endorse and that there is a linear relationship between a given construct and the total scale score attained by respondents (Andrich, 2011, Tennant et al., 2004).

In practice, items are more likely to be located along a continuum and ordered in terms of their severity (ranging from easy to difficult items to endorse). Spacing (or distance) between these items are also likely to differ according to their increments in severity (Tennant and Conaghan, 2007). The relationship between simple summated scores and the level of a construct is therefore non-linear as equal increases in units of the total score do not necessarily equate to equal sized increments (Wright and Stone, 1979) .

Various methodological approaches in measurement theory aim to address this difficulty. One approach, Rasch measurement theory (RMT), was used in this chapter to explore the possibility of achieving interval-level scaling and thus maximising the interpretability of eHIQ scores (Tennant and Conaghan, 2007). RMT operates with the assumption that the probability of a person endorsing an item is governed by the level of the construct (e.g. knowledge, ability, health status, trait, state) a person has (referred to in RMT as 'person ability') and the location of an item (termed 'item difficulty' in RMT) on a common logit scale (Wyke and Andrich, 2006). For example, in the dichotomous response case, items are located on a logit scale at the point where a person has 0.5 probability of scoring 0 or 1.

Data achieving fit to the Rasch model transforms increments of a given construct to approximate linearity, and therefore equal intervals. This enables comparisons of total scores on a given construct, in this case the domains identified as relating to the impact of using a health-related website (Wright and Stone, 1979, Andrich, 2011). Using eHIQ scores derived using the Rasch model, may improve interpretability.

Fit to the Rasch model satisfies the requirement of 'objective measurement' which essentially means that the instrument itself is not directly affected by (i.e. it is independent of) the construct being measured (Andrich, 2004a). This contrasts with CTT methods which depend on normally distributed data and variation within a specified population (Wright, 1980, Rasch, 1960, Wright and Masters, 1982).

An extension of the dichotomous Rasch model allows for polytomous Likert response options (Andrich, 1978). Likert response options are essentially expressed opinions recorded in an ordered manner. This can be challenging in measurement as responses may not represent agreement in equal increments. Furthermore, item stems are likely to be worded differently from item to item meaning simple agreement with one item (i.e. respondent selecting 'Agree') will not necessarily represent the same level of the overall construct as measured by the same response category with another item (Bond and Fox, 2013). The polytomous Rasch model was thought to be particularly useful in confirming the eHIQ response categories were performing in the way intended (i.e. if they represent successive ordering) (Andrich et al., 1997).

9.2 The Rasch model

RMT involves the probabilistic modelling of responses to an item (Andrich, 2011, Rasch, 1960). Invariance is a property of the mathematical model which enables comparisons between persons independently of the items used and vice versa (Andrich, 2004a, Andrich, 2005). It is very important to note that invariance in this instance is a property of the model and not the data (Andrich, 2004a).

Other approaches in measurement theory try to take account of increasing item difficulty within a scale. For example, Guttman scaling is desirable in measuring more or less of a construct as items are ranked in terms of increasing difficulty (Guttman, 1944, Andrich, 1982). Using a Guttman approach, eHIQ items would be ranked from 'easy' to 'difficult' to endorse. Respondents failing to endorse the fifth 'easiest' item would also be expected to fail to endorse all items ranked above the fifth easiest item (i.e. all 'harder' items in the scale). If the same respondent endorsed all items ranked below the fifth easiest item (i.e. items one through to four), their 'ability' or level on the construct being measured would be located at the fourth item.

While this approach has desirable qualities in measurement theory, it does not account for responses which contradict the expected response pattern, particularly when items difficulties or a groups abilities are close together (Andrich, 1982). To allow for these contradictions, the Rasch model follows a probabilistic Guttmann structure allowing for the possibility that a person who would be expected to endorse an 'easy' item on the lower end of the continuum and reject a 'hard' item at the higher end of the continuum to instead reject the 'easy' item while endorsing the 'hard' item (Ryan, 1983, Wright and Stone, 1979). A person's observed score in RMT is a function of the distance between the item and person location (Tennant and Conaghan, 2007). For a more detailed overview of the algorithm behind RMT and some worked examples of the model, see Appendix 9.1.

9.3 Rasch analysis

All items contributing to the eHIQ-Part 1 and Part 2 sub-scales were analysed using a priori specifications to see if data fit the model (Andrich, 2011). Analysis was performed using RUMM 2030 software. RUMM 2030 software creates class interval structures by dividing participants into the lower, middle and upper scoring groups to calculate fit statistics and graphically display the observed mean scores compared to the theoretical mean scores. The range of criteria used to examine the fit of the data were as follows:

9.3.1 Correct model to use: Likelihood Ratio test

Two derivations of the polytomous Rasch model are available: the Rating Scale Model and the Partial Credit Model (or Unrestricted Rating Scale Model). The former places constraints on the threshold parameters and requires the thresholds be equally spaced for all items while the latter is more flexible in relation to threshold parameters (Pallant and Tennant, 2007, Masters, 1982, Wright and Masters, 1982). RUMM2030 performs the Likelihood Ratio test where significance implies the Unrestricted Model should be used (Tennant and Conaghan, 2007).

9.3.2 Thresholds of response options

The polytomous graded response model was used in this analysis to take into account the probability of a person choosing any of the given response categories. The point at which there is equal probability a person will choose between two adjacent categories is referred to as the threshold (Tennant and Conaghan, 2007). It is important to note that whilst response categories are ordered in the rating scale model, this does not imply equal intervals or distances between responses categories, rather it implies equal discriminations (Andrich, 2005).

Order across response categories cannot be assumed and the hypothesis of successive ordering over the eHIQ response categories needed to be tested (Andrich, 1978). Disordered thresholds were examined through firstly looking at the threshold map output which illustrates the threshold points between response scores. At each threshold point it is equally likely that a person obtains either response score (Tennant and Conaghan, 2007). Graphical checks were made using category probability curves to confirm discrimination between response categories and evidence of person ability levels where the probability for selecting each response option was most likely (Tennant and Conaghan, 2007). Disordered thresholds were remedied by collapsing adjacent categories (i.e. response options) and retesting fit (Tennant and Conaghan, 2007).

9.3.3 Response dependency

Once the 'Rasch factor' has been accounted for, there should be no further systematic patterns among the items within the scale (i.e. all remaining variance is random noise) (Tennant and Conaghan 2007). This is referred to as the assumption of 'local independence'. Violations to this assumption can occur when responses depend on both the construct being measured and responses to other items within the scale (Embretson and Reise, 2000). For example, in a physical functioning test which includes a number of items about walking distance, responses will be linked as a respondents answer to one item will be dependent on their answer to other walking distance items. This creates bias in a test through inflating reliability and affecting parameter estimation in Rasch analysis (Tennant and Conaghan, 2007). In contrast, an item relating to the respondents lifting ability in the same physical functioning scale would be less likely to show response dependency with a walking item. Response dependency was considered to be present where residual correlations between items were >0.3 or the assumption of unidimensionality was not met.

Unidimensionality assumes an attribute can be isolated and measured using items which themselves are not influenced by other characteristics (Ryan, 1983). Unidimensionaility was assessed using Principal Components Analysis (PCA) of the standardised residuals from the person-item differences (i.e. differences between observed and actual values) (Pallant and Tennant, 2007, Smith, 2002). An independent t-test between two subsets (the most positively and the most negatively correlated items) were assessed and binomial test of proportions used to ascertain confidence intervals to take into account the sample size (Tennant and Conaghan, 2007).

9.3.4 Model fit

Person and item residuals, the differences between observed and expected scores, were examined for indications of fit to the model (Wilson, 2005). Good person and item fit are indicated by a mean of zero and a standard deviation of 1 (Pallant and Tennant, 2007). Items which did not fit the Rasch model were identified by fit residuals above 2.5 (Andrich, 1988, Pallant and Tennant, 2007). Item observed scores are compared with expected scores across class intervals using the chi-square statistic with significant values (<0.05) indicating misfit to the Rasch model (Tennant and Conaghan, 2007, Pallant and Tennant, 2007).

As the Rasch model is probabilistic, persons fit statistics allow for some minor unexpected response patterns, yet identify persons with hugely unexpected response patterns (Wilson, 2005). If all participants respond in the manner expected by the model, all person fit residuals (differences between observed and estimated scores) would be lie within the range of +/-2.5. Those outside this range are referred to as ‘misfitting persons’ and may require further investigating to see if there is any reason they do not respond in the same way as others completing the questionnaire. Where a small number of persons with high fit residuals (>2.5) were evident, they were removed to see if they had a serious impact on the fit statistics (Tennant and Conaghan, 2007, Linacre, 2010, Curtis, 2004). While removing severely ‘misfitting persons’ from the analysis may have improved internal validity, care was taken to identify any reasons behind why such respondents data did not fit the model, for example, whether persons were all from a particular condition or age group (Tennant and Conaghan, 2007).

The chi-square interaction statistic, which is calculated by adding the chi-square values for the individual scale items, was used to assess overall fit to the model. Significant values indicated hierarchical ordering varied and the probability of invariance (of the model) was compromised (Tennant and Conaghan, 2007, Pallant and Tennant, 2007). The Bonferroni adjusted level was used at the 0.05 significance level (Bland and Altman, 1995).

9.3.5 Construct validity

In order for a scale to be considered valid, it should satisfy the requirement of invariance and have the same meaning across groups. Differential item functioning (DIF) occurs where there is bias in the manner an item performs across population groups (for example, across age groups or gender), despite otherwise having the same score on the underlying construct (Tennant and Conaghan, 2007). In other words, DIF occurs where an item is interpreted differently within a specific population group (i.e. the group has a different probability of endorsing the item than another group) even though they have the same level of an attribute. For example, women may be expected to achieve a higher score on a particular item, however, men and women should still have the same probability of endorsing the item if they have the same level on a given construct. Therefore, where DIF is present, the probability of a person with certain ‘ability’ endorsing an item is different across groups, compromising the instruments measurement properties.

Expected values according to the model were examined graphically using the Item Characteristic Curve (ICC). The presence of ‘uniform’ DIF was indicated by different expected values across groups, but parallel ICC slopes. ‘Non-uniform’ DIF was indicated by the same expected values across groups but non-parallel slopes. Non-uniform DIF was also evident when the expected values and their slopes were different (Andrich, 2004b). Equivalence across sub-groups was tested statistically across the models class intervals using ANOVA (Andrich, 2004b). Values below the Bonferroni adjusted significance level indicated significant DIF (Bland and Altman, 1995). Where appropriate items were split for DIF to allow different item parameters for the relevant groups.

9.3.6 Internal reliability

Internal consistency in CTT is calculated using the Cronbach’s alpha statistic which tests the interrelationship between the raw scores (Nunnally and Bernstein, 1994). Person Separation Index (PSI), the equivalent to the Cronbach’s alpha statistic, was used in this analysis (Fisher, 1992). The PSI is calculated using non-linear transformations of raw scores derived from the item and person parameters and takes into account the variances of person estimates relative to the error variance of each person (Andrich, 1982). A PSI value of 0.7 or above was considered an acceptable level of internal reliability (Bond and Fox, 2013).

9.3.7 Targeting

Appropriate targeting was analysed by looking at the mean location of the person's scores. Appropriately targeted scales (i.e. overall, the items are not too easy or not too difficult) would be expected to have a mean location of approximately zero (Pallant and Tennant, 2007). Scales with 'too easy' items would display high ceiling effects where most respondents endorse all items, whilst scales with items that are 'too difficult' would display floor effects where respondents endorse very few items.

The graphical output from RUMM 2030 was also used to check the spread of items across scores. The person-item distribution map is used to show how the items and person responses are distributed along a single logit scale. The upper part of the graphs represent persons and their respective ability levels while the lower part gives item locations on the common continuum and their distributions. As the (polytomous) rating scale model was used, the person-item threshold distribution was examined as it is more representative of the data.

9.3.8 Scoring

Person locations scores were exported and used to conduct parametric analyses in this chapter. RUMM 2030 software converts locations to the original score range of the respective scale. Scores are not converted to a standard 0-100 range as this inflates the scoring error. Person location scores were transformed to raw scores within RUMM 2030 software and exported to transformation tables.

9.4 Results

9.5 Equal interval scaling (eHIQ Part 1)

For ease of interpretation sub-scales are referred to throughout the text as:

Scale 1.1: *Attitudes towards online health information* (5 items)

Scale 1.2: *Attitudes towards sharing health experiences online* (6 items)

Responses (n=170) for the eleven eHIQ-Part 1 items were entered into RUMM 2030. Sub-scales were analysed independently. Extreme scores (i.e. where participants had either endorsed all items or did not endorse any item) provide no new information to the Rasch model. Three participants with extreme scores in Scale 1.1 were therefore removed by the software, leaving 167 for analysis. Twelve participants in Scale 1.2 had extreme scores or invalid records, leaving 158 for analysis. Significant Likelihood Ratio tests indicated the Unrestricted Model should be used and not the Rating Scale model.

Class interval structures (i.e how participants are divided up into the lower, middle and upper scoring groups) were examined for both scales to ensure the sample was distributed adequately. This is important as class intervals are used to calculate fit statistics and graphically display the observed mean scores compared to the theoretical mean scores. RUMM2030 specified three class intervals which ranged from 54-58 participants in each interval for Scale 1.1 and 42-68 for Scale 1.2. On entry, the initial Scale 1.1 mean fit residual was -0.18 (SD 1.2) for items and -0.53 (SD 1.14) for persons. For Scale 1.2, the mean fit residual was -0.11 (SD 1.15) for items and -0.63 (SD 1.52) for persons (see Table 9-1, Models 1.1.1 and 1.2.1). Although these residuals demonstrated reasonable fit between the data and the model, some degree of misfit was present which was subsequently investigated.

Thresholds were inspected to see if disordered item categories were affecting model fit. Thresholds were ordered for Scale 1.2; however, disordered thresholds were evident for one item (Item 7) in Scale 1.1. The second and third response categories ('Disagree' and 'Neither Agree nor Disagree') were collapsed resulting in ordered thresholds and all category frequencies for items in each scale are given in Table 9-2 and Table 9-3. Threshold maps for both scales showed evidence of ordered thresholds (see Figure 9-1 and Figure 9-2). Ordering of Item 7 thresholds resulted in a slight improvement in overall fit for Scale 1.1: Item fit: -0.08 (SD 0.94), Person fit: -0.55 (SD 1.20) (see Table 9-1, Model 1.1.2). The two non-significant Chi-square values indicated hierarchical ordering of items (Scale 1 .1: x^2 =15.63, p= 0.11 and Scale 1.2: x^2 =25.36, p= 0.01). The assumption of local independence was met for both scales and both scales demonstrated good internal reliability (PSI ≥ 0.76).

Table 9-1 Model fit statistics (eHIQ- Part 1)

Model	**Description**	**Item fit residual** [a] *(SD)*	**Person fit residual** [a] *(SD)*	χ^2 *p-value*	**PSI** [b]	**Unidimensionality** [c] *(CI)*
Scale 1.1						
1.1.1	Scale 1.1, disordered thresholds	-0.18 (1.2)	-0.53 (1.14)	0.04	0.77	0.08 (0.050/0.115)
1.1.2	Scale 1.1, ordered thresholds	*-0.08* *(0.94)*	*-0.55* *(1.20)*	0.11	0.76	*0.08* *(0.050/0.115)*
1.1.3	Scale 1.1 with 21 persons deleted	*0.25* *(0.87)*	*-0.33* *(0.98)*	0.24	0.77	*0.09* *(0.052/0.122)*
Final model 1.1.4	Scale 1.1, no persons deleted, Item 7 spilt for DIF (Sex)	*-0.07* *(0.81)*	*-0.53* *(1.19)*	0.13	0.76	*Not available following item split*
Scale 1.2						
1.2.1	Scale 1.2, ordered thresholds	-0.11 *(1.15)*	-0.63 (1.52)	0.01	0.84	0.03 *(0.000/0.062)*
Final model 1.2.2	Scale 1.2 with 19 persons deleted	0.25 (1.08)	-0.31 (1.06)	0.0245	0.85	0.07 (0.032/0.102)

[a] Item and person fit residuals should approximate 0 with a standard deviation (SD) of 1.

[b] Person Separation Index (PSI) of 0.70 is acceptable internal consistency and 0.80 is good.

[c] Confidence interval for proportion of significant *t*-tests, lower boundary ≤0.05 for unidimensionality.

Table 9-2 Category frequencies for Scale 1.1, Model 1.1.2

Item	n	Response				
		1	2	3	4	5
2. The internet is a reliable resource to help me understand what a doctor tells me.	167	2	12	54	83	16
5. The internet can help the public to know what it is like to live with a health problem.	167	0	5	18	119	25
6. The internet can be useful to help people decide if their symptoms are important enough to go to see a doctor.	167	3	10	36	106	12
7. I would use the internet if I needed help to make a decision about my health (for example, whether I should see a doctor, take medication or seek other types of treatment).*	167	10	59	83	15	-
8. I would use the internet to check that the doctor is giving me appropriate advice.	166	6	46	45	57	12

* Item 7 (Rescored: 01123- ‘Disagree’ and ‘Neither agree nor disagree’) collapsed due to disordered thresholds.

Table 9-3 Category frequencies for Scale 1.2, Model 1.2.1

Item	n	Response				
		1	2	3	4	5
9. The internet is a good way of finding other people who are experiencing similar health problems.	158	0	9	28	93	28
10. It can be helpful to see other people’s health-related experiences on the internet.	158	0	13	30	91	24
11. The internet is useful if you don’t want to tell people around you (for example, your family or people at work) how you feel.	158	2	17	46	74	19
12. It can be reassuring to know that I can access health-related websites at any time of the day or night.	158	1	7	29	88	33
13. The internet is a good way of finding other people who are facing health-related decisions I may also face.	158	0	12	32	93	21
14. Looking at health websites reassures me that I am not alone with my health concerns.	158	1	11	43	87	16

Figure 9-1 Threshold plot for Scale 1.1, Model 1.1.2

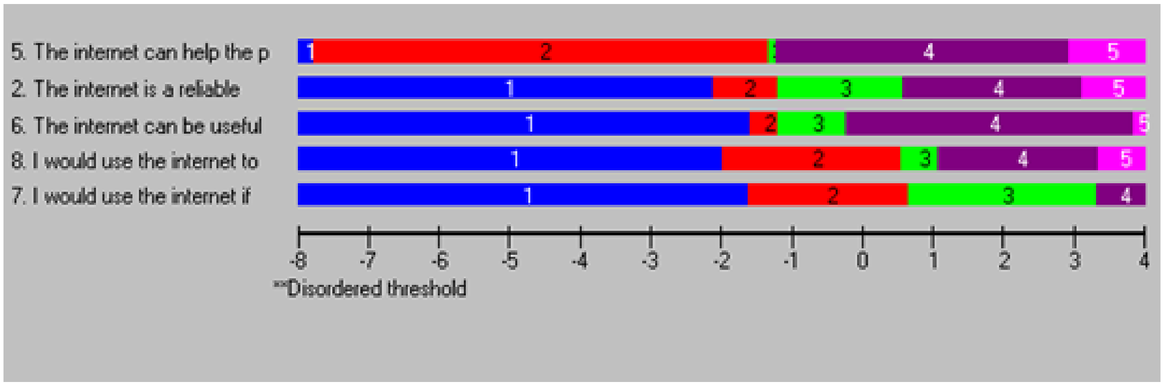
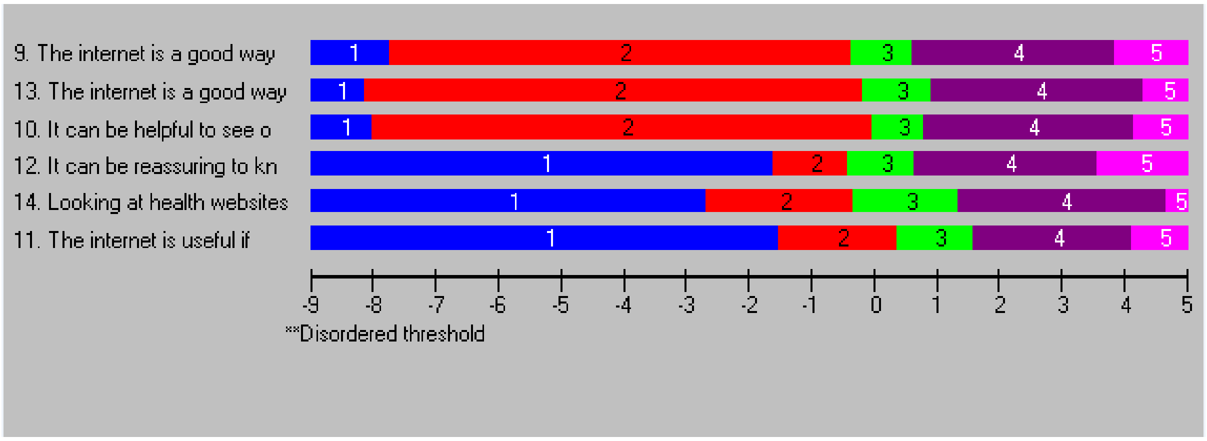


Figure 9-2 Threshold plot for Scale 1.2, Model 1.2.1



9.5.1 Model fit (eHIQ-Part 1)

Following the reordering of Item 7, individual item fit was examined (see Tables 9-5 and 9-6). All item fit residuals fell within the expected +/-2.5 values. Chi-square significance for all items was greater than their Bonferroni adjusted value, indicating no significant deviation from the model.

Table 9-4 Item fit for Scale 1.1, Model 1.1.2

Item	Location	FitResid	ChiSq	Prob
2	0.095	1.349	7.108	0.029
5	-1.848	0.282	1.408	0.495
6	0.214	-1.005	3.707	0.157
7	0.789	-0.766	1.577	0.455
8	0.750	-0.246	1.833	0.400

Bonferroni adjustment for 5 items at 0.01 level=0.002

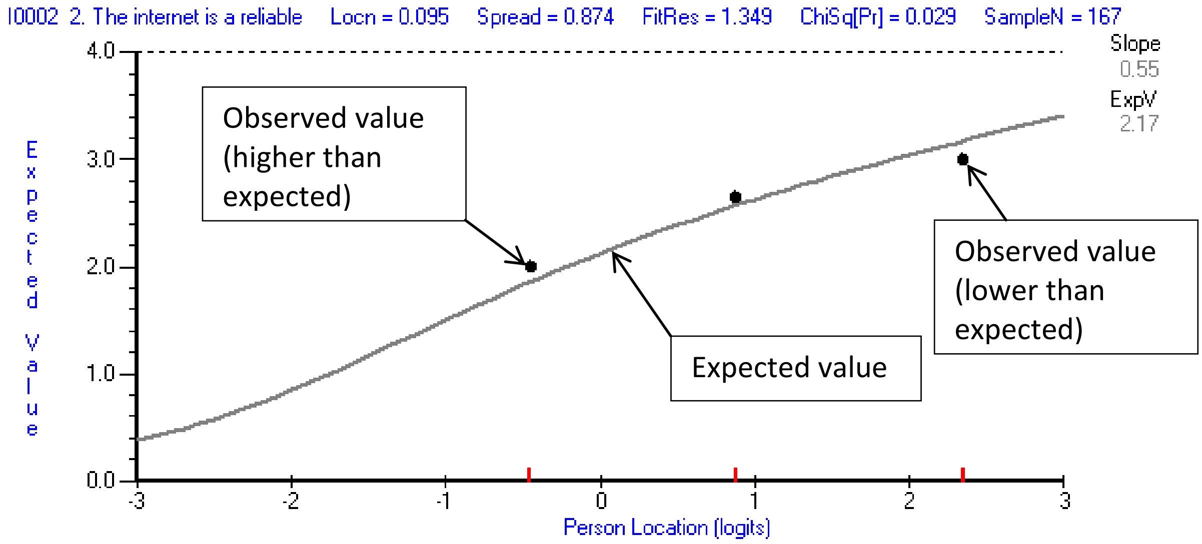
Table 9-5 Item fit for Scale 1.2, Model 1.2.1

Item	Location	FitResid	ChiSq	Prob
9	-0.908	-0.771	4.389	0.111
10	-0.770	-1.306	5.136	0.077
11	1.148	1.477	4.613	0.100
12	0.549	1.000	5.917	0.052
13	-0.772	-1.092	4.566	0.102
14	0.752	0.032	0.741	0.690

Bonferroni adjustment for 6 items at 0.01 level=0.0017

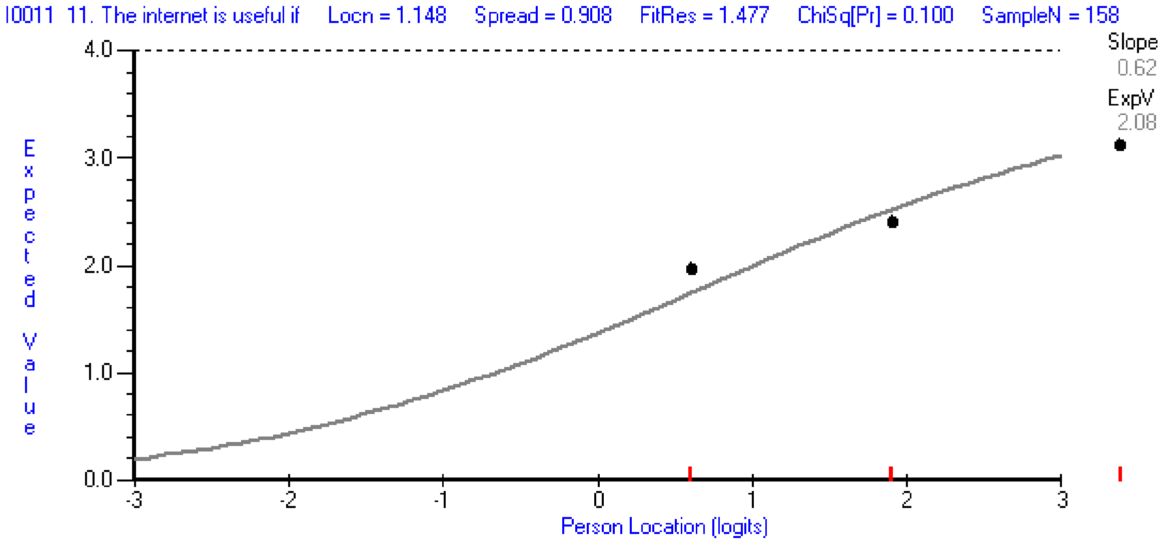
Items 2 and 11 were the worst fitting items for Scale 1.1 and Scale 1.2 respectively. Item 2 (item fit 1.35, p= 0.029) indicated low levels of item discrimination. The ICC curve which plots observed means against the curve of expected means for each class interval confirmed low discrimination. This meant that the probability of endorsing easier items was higher than the expected in the model for Item 2 whilst the observed probability for the more difficult items was lower than expected (see Figure 9-3). Item 11 also indicated low levels of discrimination (Item fit 1.48, p=0.10). Observed values were higher than expected for items located lower on the difficulty continuum, but lower than expected as the item difficulty increased (see Figure 9-4). Systematic deviations from the model however were not statistically evident with either item. These items therefore do not fit the Rasch model perfectly, but their residuals are sufficiently close to zero that the Chi-square did not indicate misfit.

Figure 9-3 Item 2 Item Characteristic Curve (Scale 1.1, Model 1.1.2)



Item 2: The internet is a reliable resource to help me understand what a doctor tells me.

Figure 9-4 Item 11 Item Characteristic Curve (Scale 1.2, Model 1.2.1)



11. The internet is useful if you don’t want to tell people around you (for example, your family or people at work) how you feel.

Individual person fit statistics for Scale 1.1 showed that 21 persons had residuals outside the acceptable range (>2.5). On their removal, the overall model fit statistics improved slightly (Item fit: 0.25 (SD 0.87), Person fit: -0.33 (SD 0.98)) (see Table 9-1, Model 1.1.3). It was decided however to retain the persons as improvement was marginal. Individual person fit statistics for Scale 1.2 showed that 19 respondents had residuals outside the acceptable range (>2.5). On their removal the overall model fit statistics improved (Item fit: 0.25 (SD 1.08), Person fit: -0.31 (SD 1.06)) (see Table 9-1, Model 1.2.2). It was decided to remove these persons to improve model fit.

Presence of DIF was tested for both scales. Item 7 (Scale 1.1, Model 1.1.2) showed evidence of uniform DIF for Sex (i.e. males and females answered the item slightly differently despite having the same level of the construct), while no DIF was evident for any item in Scale 1.2, Model 1.2.2. Item 7 was split for Sex improving the model fit marginally (see Table 9-1, Model 1.1.4). The two scales showed reasonably good fit to the Rasch model (Scale 1.1- Item fit: -0.07 (SD 0.81), Person fit: -0.53 (SD 1.19) and Scale 1.2- Item fit: 0.25 (SD 1.08), Person fit: -0.31 (SD 1.06)). PCA analysis showed both scales were unidimensional (see Table 9-1) and all scales had non-significant Chi-square values indicating items were largely hierarchical.

9.5.2 Scale targeting (eHIQ- Part 1)

The average mean person location value for Scale 1.1 was 0.80 (SD 1.51) suggesting that the scale on the whole was reasonably well-targeted. People on average were positive towards using online health information. The average mean location for Scale 1.2 was 2.26 (SD 1.87) suggesting that these items, on average, were quite easy for respondents to endorse.

Rasch person-item threshold distribution maps for Scales 1.1 and 1.2 are shown in Figure 9-5 and Figure 9-6. The items for both scales are reasonably well spaced out with some location points having more than one item (or item threshold) to measure their locations. There are some gaps between items in both scales which indicate measurement would be more accurate if improvements were made to item targeting. The maps show the distributions of people and items across the continuum of each construct to be relatively normal but distinct ceiling effects for Scale 1.2.

Respondents with the lowest scores have the most negative views of each construct and items at the highest part of the scale can only be endorsed by those with the highest level of each construct. Despite some floor and ceiling effects, a good range of measurement is evident as items are spread across the whole spectrum of difficulty. This indicates some items were very easy to endorse and some items were difficult for the sample.

Figure 9-5 Person-Item threshold distributions for Scale 1.1, Model 1.1.4

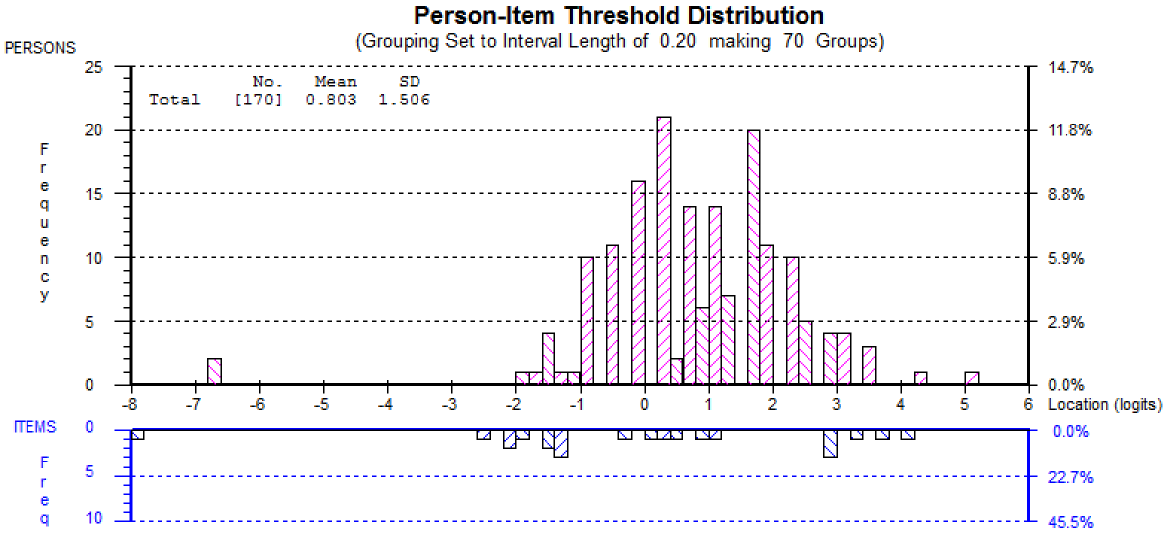
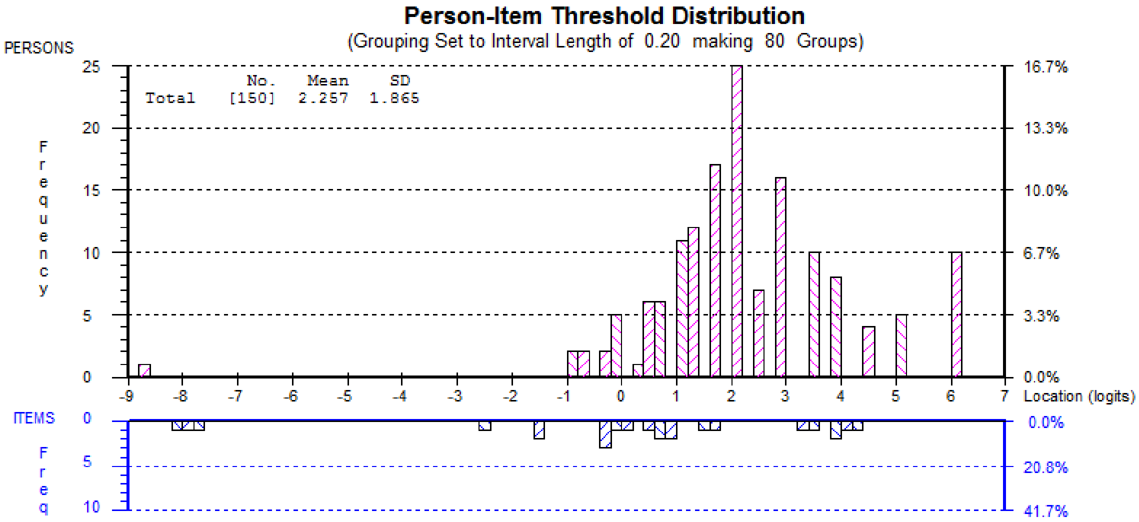


Figure 9-6 Person-Item threshold distributions for Scale 1.2, Model 1.2.2



9.5.3 Scoring (eHIQ-Part 1)

Person location values were transformed to raw scores using RUMM2030 software. Raw scores were plotted against location values indicating raw scores are not linear across the full range of each scale. Figure 9-7 and Figure 9-8 indicate a small degree in movement towards the lower ends of each scale will result in a substantial difference in score compared to movement in the middle range of the constructs. Measurement may therefore be more precise towards the central level of the construct. See Appendix 9.2 for raw score conversion tables.

Figure 9-7 Scale 1.1 Location value from RUMM2030 against raw score

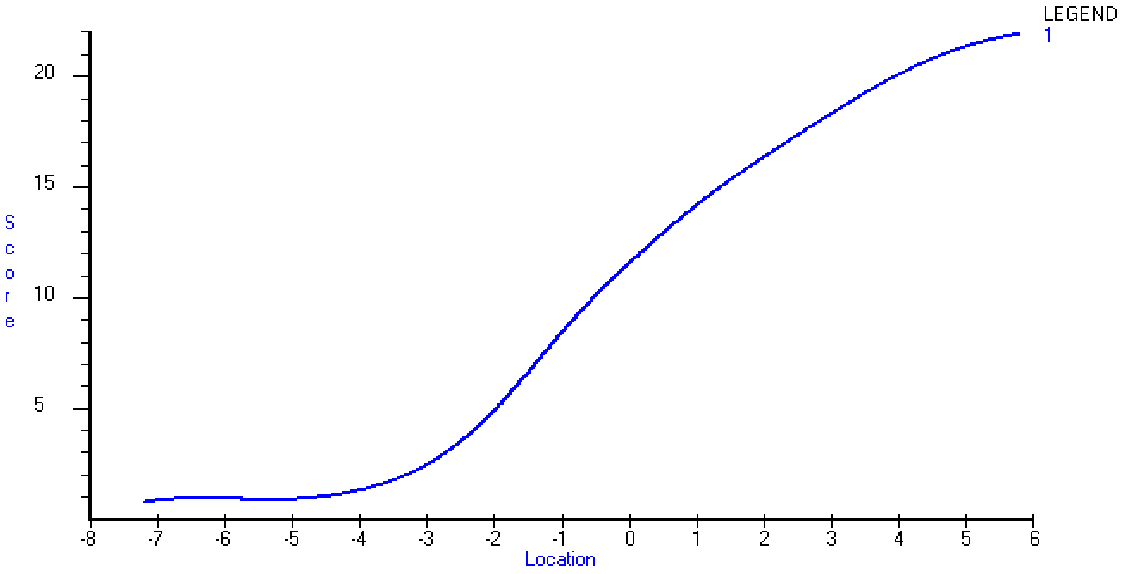
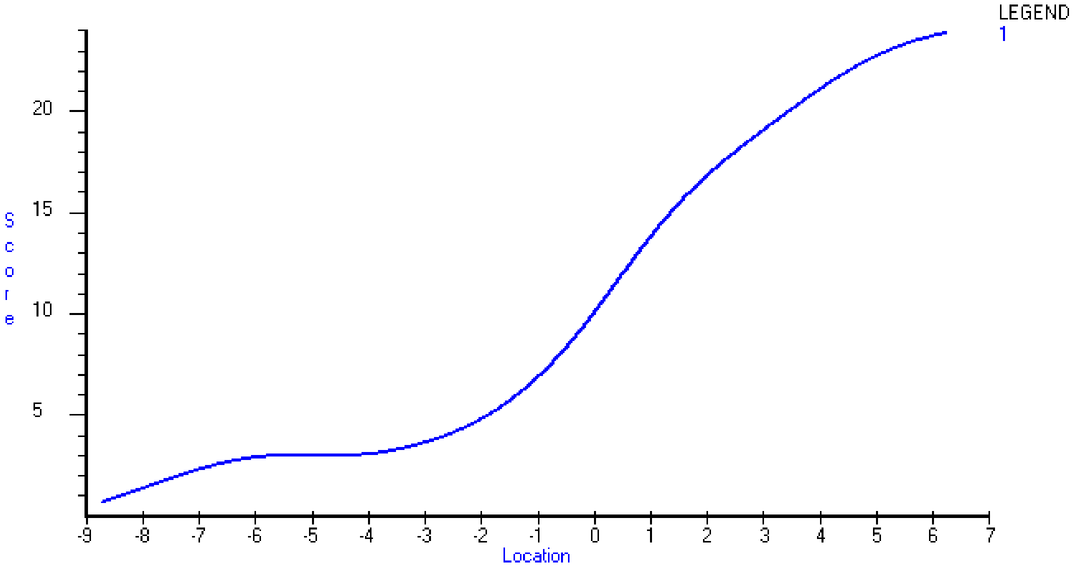


Figure 9-8 Scale 1.2 Location value from RUMM2030 against raw score



9.6 Equal interval scaling (eHIQ Part 2)

For ease of interpretation sub-scales are referred to throughout the text as:

Scale 2.1: *Confidence and identification* (9 items)

Scale 2.2: *Information and presentation* (8 items)

Scale 2.3: *Understanding and motivation* (9 items)

Responses (n=168) for the 26 items in the eHIQ-Part 2 items were entered into RUMM 2030. Sub-scales were analysed independently. Two participants from Scales 2 .1 and Scale 2.2 had extreme scores leaving 166 for analysis and while Scale 2.3 retained all participants for analysis. Significant Likelihood Ratio tests indicated the Unrestricted Model should be used and not the Rating Scale model.

Class interval structures were examined for all scales to ensure the sample was adequately distributed. RUMM2030 specified three class intervals which ranged from 51-58 participants in each interval for Scale 2.1, 47-64 for Scale 2.2 and 55-56 for Scale 2.3. On entry, the initial mean fit residuals for Scale 2.1 were -0.25 (SD 1.37) for items and -0.73 (SD 1.81) for persons (see Table 9-6, Model 2.1.1). For Scale 2.2, the mean fit residual was -0.60 (SD 1.42) for items and -0.59 (SD 1.28) for persons and for Scale 2.3 the mean fit residual was -0.01 (SD 0.80) for items and -0.42 (SD 1.25) for persons (see Table 9-6, models 2.2.1 and 2.3.1). All scales had non-significant Chi-square values however some degree of misfit was evident according to their fit residuals.

Thresholds, the point at which there is equal probability a person will choose between two adjacent categories, were inspected for items within each scale to see if respondents interpreted response options in a successive order. Disordered item categories (or incidents where participants do not treat response categories successively) can were affect model fit. Thresholds were ordered for Scale 2.1 while disordered thresholds were disordered were evident in Scale 2.2 (Items 17, 31, 32) and Scale 2.3 (Item 10). Disordered thresholds were ordered through collapsing two adjacent categories for each item (See Table 9-7, Table 9-8 and Table 9-9 for reordering categories where applicable). This effectively gives the same score for the two adjacent collapsed categories.

The best fitting solutions following ordering of thresholds showed some improvement in overall fit residuals for Scale 2.2 (Item fit: -0.56 (SD 1.20), Person fit: -0.57 (SD 1.31)) and Scale 2.3 (Item fit: -0.21 (SD 0.70), Person fit-0.45 (SD 1.21)) (see Table 9-6, Models 2.2.2 and 2.3.2). Threshold maps for both scales showed evidence of ordered thresholds (see Figure 9-9, Figure 9-10 and Figure 9-11). The three non-significant Chi-square values indicated hierarchical ordering of items (Scale 2.1: x^2 =25.26, p= 0.118, Scale 2.2: x^2 =17.85, p= 0.33, Scale 2.3: x^2 =12.41, p = 0.83). All scales met the requirement of local independence and demonstrated good internal reliability (PSI ≥ 0.85).

Table 9-6 Model fit statistics (eHIQ- Part 2)

Model	**Description**	**Item fit residual** [a] *(SD)*	**Person fit residual** [a] *(SD)*	x^2 *p-value*	**PSI** [b]	**Unidimensionality** [c] *(CI)*
Scale 2.1						
2.1.1	Scale 1.1, ordered thresholds	-0.25 (1.37)	-0.73 (1.81)	0.12	0.91	0.089 (0.056/0.122)
Final model: 2.1.2	Scale 2.1 with 22 persons deleted	0.14 (1.38)	-0.31 (1.14)	0.04	0.90	0.089 (0.054/0.124)
Scale 2.2						
2.2.1	Scale 2.2, disordered thresholds	-0.60 (1.42)	-0.59 (1.28)	0.54	0.85	0.089 (0.056/0.122)
Final model: 2.2.2	Scale 2.2, ordered thresholds	-0.56 (1.20)	-0.57 (1.31)	0.33	0.85	0.083 (0.05/0.116)
Scale 2.3						
2.3.1	Scale 2.3, disordered thresholds	-0.01 (0.80)	-0.42 (1.25)	0.52	0.90	0.089 (0.056/0.122)
Final model: 2.3.2	Scale 2.3, ordered thresholds	-0.21 (0.70)	-0.45 (1.21)	0.83	0.89	*0.065 (0.033/0.098)*

[a] Item and person fit residuals should approximate 0 with a standard deviation (SD) of 1.

[b] Person Separation Index (PSI) of .70 is acceptable and .80 is good.

[c] Confidence interval for proportion of significant *t*-tests, lower boundary ≤0.05 for unidimensionality.

Table 9-7 Category frequencies for Scale 2.1, Model 2.1.1

Item	n	Response				
		1	2	3	4	5
13. The website prepares me for what might happen to my health.	166	4	26	45	84	7
15. The people who have contributed to the website understand what is important to me.	165	5	16	64	72	8
19. I feel I have a sense of solidarity with other people using the website.	165	11	32	65	46	11
21. I can identify with other people using the website.	166	4	17	65	69	11
24. I value the advice given on the website.	165	5	7	39	103	11
25. The website gives me confidence that I am able to manage my health.	166	5	16	63	76	6
26. I feel I have a lot in common with other people using the website.	166	6	24	59	66	11
27. The website gives me the confidence to explain my health concerns to others.	166	6	20	70	66	4
30. The website makes me more confident to discuss my health with the people around me (for example, my family or people at work).	166	7	31	71	50	7

Table 9-8 Category frequencies for Scale 2.2, Model 2.2.2

Item	n	Response				
		1	2	3	4	5
3. The information on the website left me feeling confused.	165	3	7	32	92	31
6. The website provides a wide range of information.	166	1	4	14	112	35
7. The language on the website made it easy to understand.	166	1	2	11	121	31
11. I can easily understand the information on the website.	165	1	2	11	110	41
17. I trust the information on the website.*	166	4	45	101	16	-
31. Photographs and other images were used appropriately on the website.**	166	10	45	100	11	-
32. I found the images on the website distressing.***	166	1	39	85	41	-
34. The website is easy to use.	166	1	3	20	115	27

Items 17, 31 and 32 have four categories after reordering to account for disordered thresholds. *Item 17 (Rescored: 01123- ‘Disagree’ and ‘Neither agree nor disagree’ collapsed), **Item 31 (Rescored: 00123-

‘Strongly disagree’ and ‘Disagree’ collapsed) and ***Item 32 (Reversed and rescored: 01123 ‘Agree and ‘Neither agree nor disagree’ collapsed)

Table 9-9 Category frequencies for Scale 2.3, Model 2.3.2

Item	n	Response				
		1	2	3	4	5
1. The website encourages me to take actions that could be beneficial to my health.	167	3	13	54	78	19
2. The website has a positive outlook.	168	0	2	34	105	27
4. The website includes useful tips on how to make life better.	168	2	12	38	99	17
9. I feel more inclined to look after myself after visiting the website.	168	8	34	74	50	2
10. I have learnt something new from the website.*	168	12	30	115	11	-
18. I would consult the website if I had to make a decision about my health.	168	9	27	46	78	8
22. On the whole, I find the website reassuring.	168	5	8	67	79	9
28. The website helps me to have a better understanding of my personal health.	168	6	12	57	86	7
29. The website encourages me to play a more active role in my healthcare.	168	7	20	63	69	9

Item 10 has four categories after reordering to account for disordered thresholds (Rescored: 01123- ‘Disagree’ and ‘Neither agree nor disagree’ collapsed).

Figure 9-9 Threshold plot for Scale 2.1, Model 2.1.1

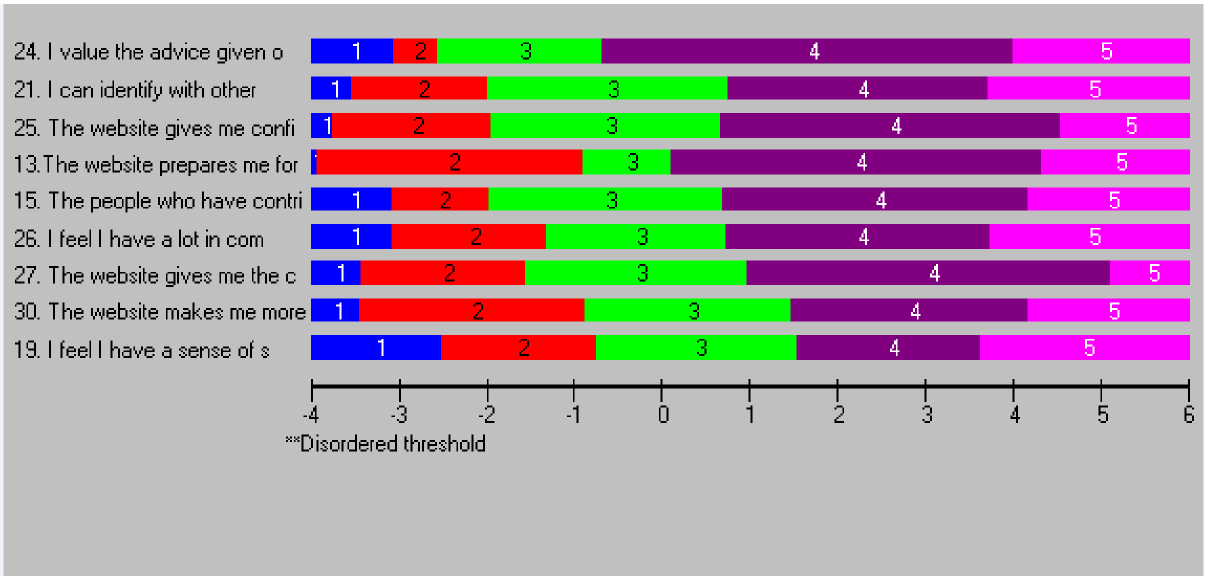


Figure 9-10 Threshold plot for Scale 2.2, Model 2.2.2

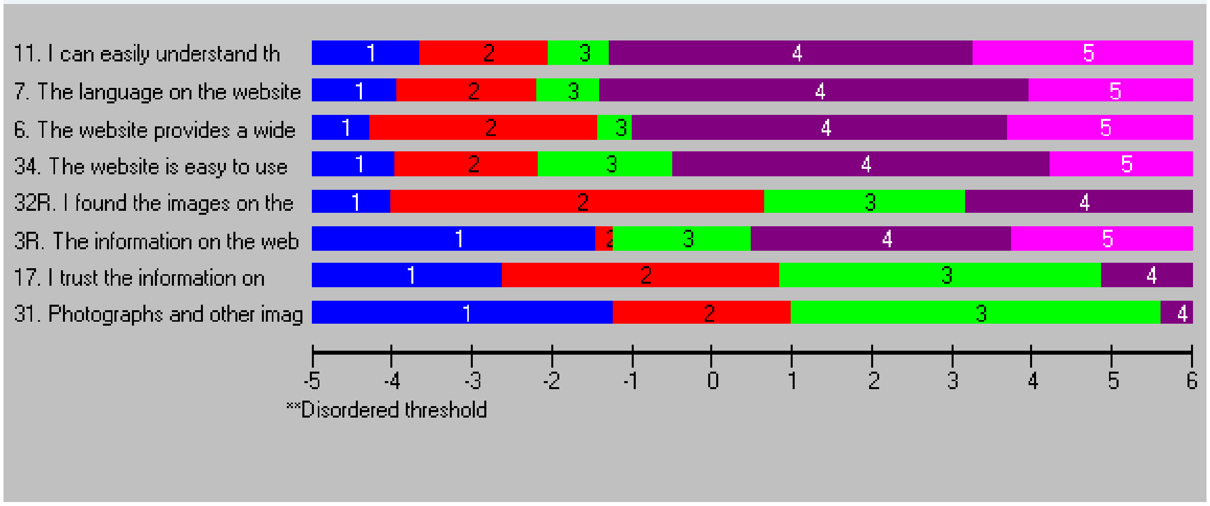
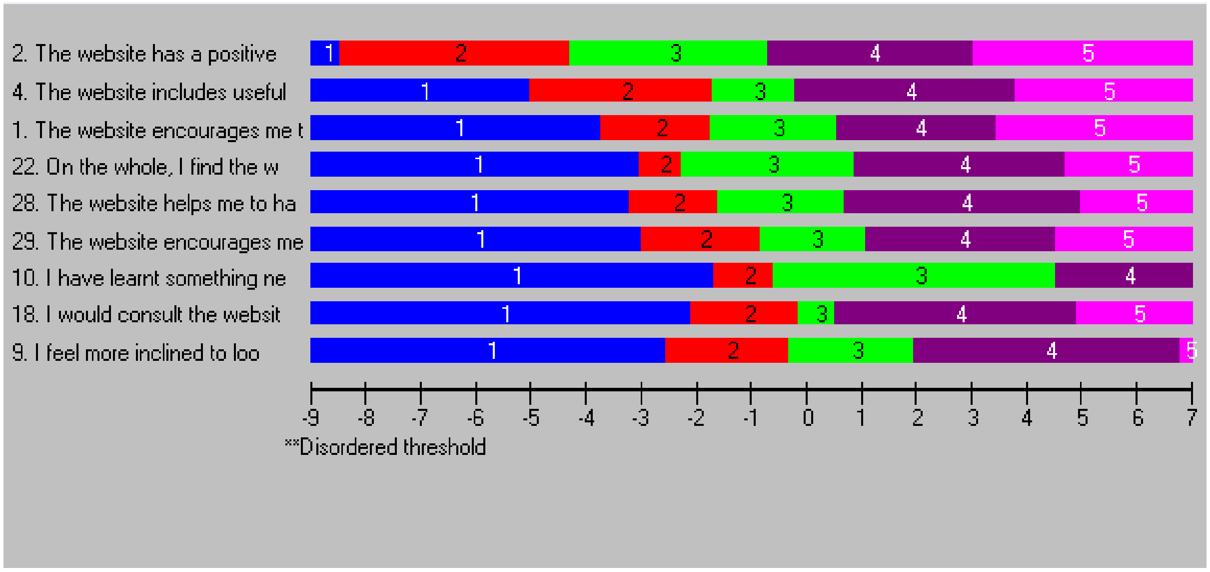


Figure 9-11 Threshold plot for Scale 2.3, Model 2.3.2



9.6.1 Model fit (eHIQ- Part 2)

Following the reordering of disordered thresholds, individual item fit was examined (see Table 9-10, Table 9-11 and Table 9-12). All item fit residuals fell within the expected +/-2.5 values. Chi-square significance for all items was greater than the Bonferroni adjusted value, indicating no significant deviation from the model.

Table 9-10 Item fit for Scale 2.1, Model 2.1.1

Item	Location	FitResid	ChiSq	Prob
13	-0.099	2.160	4.842	0.089
15	-0.044	1.259	7.323	0.026
19	0.473	-2.264	3.566	0.168
21	-0.262	-0.606	0.308	0.857
24	-0.575	-1.472	1.149	0.563
25	-0.117	-0.852	3.511	0.173
26	0.020	0.629	2.967	0.227
27	0.276	-0.535	1.277	0.528
30	0.329	-0.590	0.313	0.855

Bonferroni adjustment for 9 items at 0.01 level= 0.0011

Table 9-11 Item fit for Scale 2.2, Model 2.2.2

Item	Location	FitResid	ChiSq	Prob
3	0.396	0.423	1.794	0.408
6	-0.747	-0.152	2.567	0.277
7	-0.893	-2.005	1.747	0.417
11	-0.925	-1.383	4.024	0.134
17	1.035	-0.676	0.643	0.725
31	1.794	-0.713	3.296	0.192
32	-0.065	1.666	2.774	0.250
34	-0.595	-1.611	1.001	0.606

Bonferroni adjustment for 8 items at 0.01 level= 0.0013

Table 9-12 Item fit for Scale 2.3, Model 2.3.2

Item	Location	FitResid	ChiSq	Prob
1	-0.363	-0.382	0.023	0.989
2	-2.608	0.490	0.233	0.890
4	-0.776	-1.099	1.807	0.405
9	1.471	0.556	1.471	0.479
10	0.763	-0.055	3.630	0.163
18	0.789	0.356	1.314	0.518
22	0.071	-0.578	1.566	0.457
28	0.212	-1.395	2.254	0.324
29	0.441	0.218	0.108	0.947

Bonferroni adjustment for 9 items at 0.01 level= 0.0011

Item 19 (Fit: -2.26, p=0.17), Item 7 (Fit: -2.01, p=0.41) and Item 28 (Fit: -1.40, p=0.32) were the worst fitting items for Scales 2.1, 2.2 and 2.3 respectively. All residuals were negative indicating some misfit, however, all were within expected range and were therefore not considered to be redundant items. Item 15 also had a low Chi-square probability (p=0.03), but was non-significant at the Bonferroni level. All ICC curves showed low discrimination of items, however, all were considered acceptable with no statistically significant systematic deviations from the model (see Figure 9-12, Figure 9-13 and Figure 9-14).

Figure 9-12 Scale 2.1, Item 19 Item Characteristic Curve

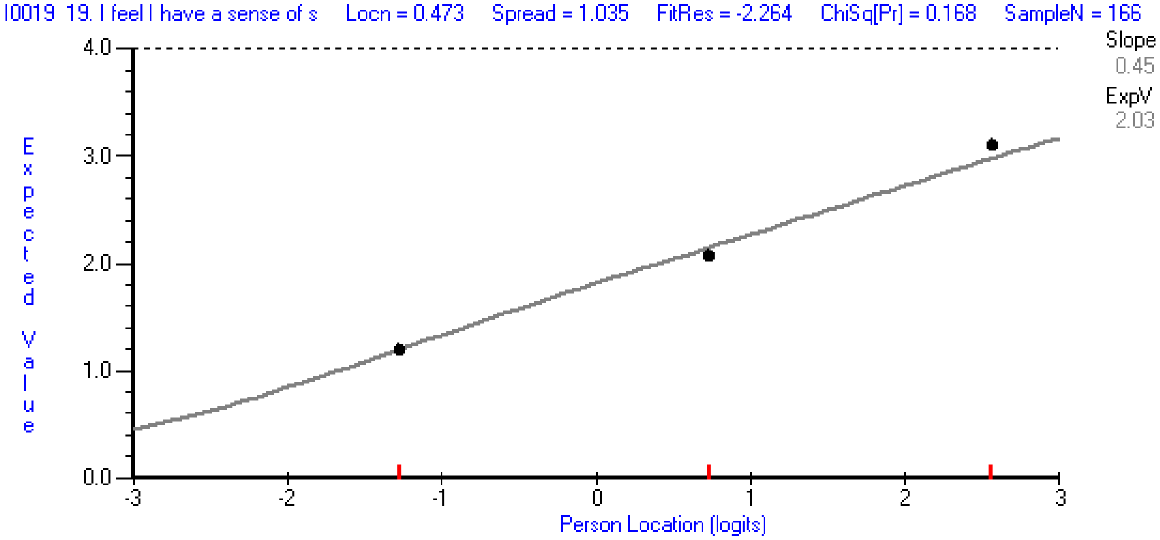


Figure 9-13 Scale 2.2, Item 7 Item Characteristic Curve

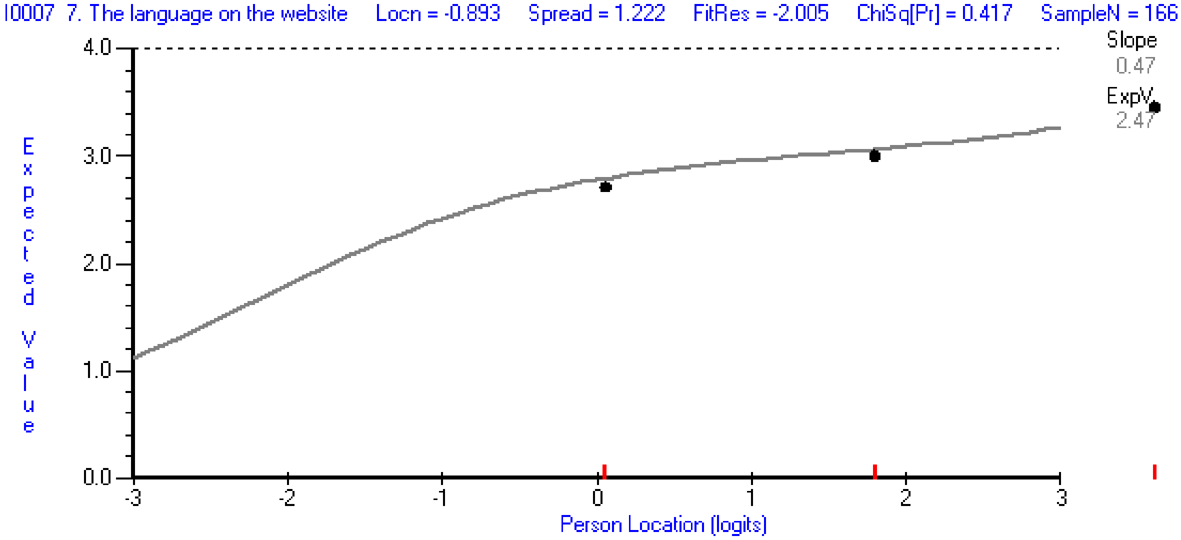
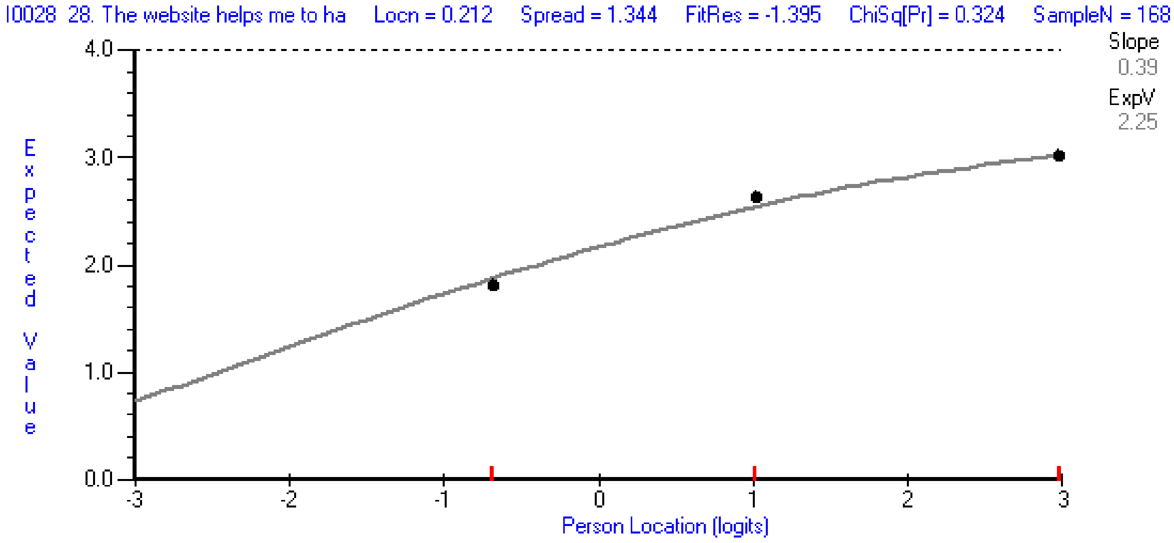


Figure 9-14 Scale 2.3, Item 28 Item Characteristic Curve



Individual person fit statistics for Scale 2.1 showed that 22 persons had residuals outside the acceptable range (>2.5). On their removal, the overall model fit statistics improved slightly Item fit: 0.14 (SD 1.38), person fit -0.31 (SD 1.13). It was decided to remove these persons as fit was improved. Individual person fit statistics for Scale 2.2 or 2.3 showed no improvements when removing high fit residuals so persons were retained for theses scales.

Items 19 and 25 (Scale 2.1) showed evidence of uniform DIF for Condition, however these were likely due to sample size. Post-hoc analysis of covariance of Item 19 mean scores showed significant differences between conditions: Healthy Eating (n=48), MND (n=32) and Cancer survivors (n=26). Post-hoc tests for Item 25 did not reveal significant differences between groups at the 0.05 significance level. No action was therefore taken to split the items for DIF. No DIF was evident for items in Scales 2 .2 and 2.3. PCA analysis showed both scales were unidimensional (see Table 9-6).

9.6.2 Scale targeting (eHIQ-Part 2)

The mean person location value for Scale 2.1 was 0.46 (SD 1.78) suggesting that the scale on the whole was reasonably well-targeted. People on average had a high score for Scale 2.1 *Confidence and identification* with the websites viewed but were largely centred around the mean. The average mean location for Scale 2.2 *Information and presentation* was 1.92 (SD 1.89) suggesting that these items, on average, were quite easy for respondents to endorse. Scale 2.3 had a mean person location value of 1.0 (SD 1.80) suggesting a reasonably high level of *Understanding and motivation* in relation to the websites viewed.

Rasch person-item threshold distribution maps for Scales 2.1, 2.2 and 2.3 are shown in Figure 9-15, Figure 9-16 and Figure 9-17. The items for all scales are reasonably well spaced out with some location points having more than one item or item threshold to measure its locations. Scale 2.2 shows quite a big gap between locations 1.0- 3.0 indicating participants at this level of the construct may be measured with less precision. Items are however spread across the whole spectrum of difficulty indicating some items were very easy to endorse and some items were difficult for the sample. Scale 2.3 shows reasonably good coverage of items on the difficulty continuum.

Figure 9-15 Person-Item threshold distributions for Scale 2.1, Model 2.1.2

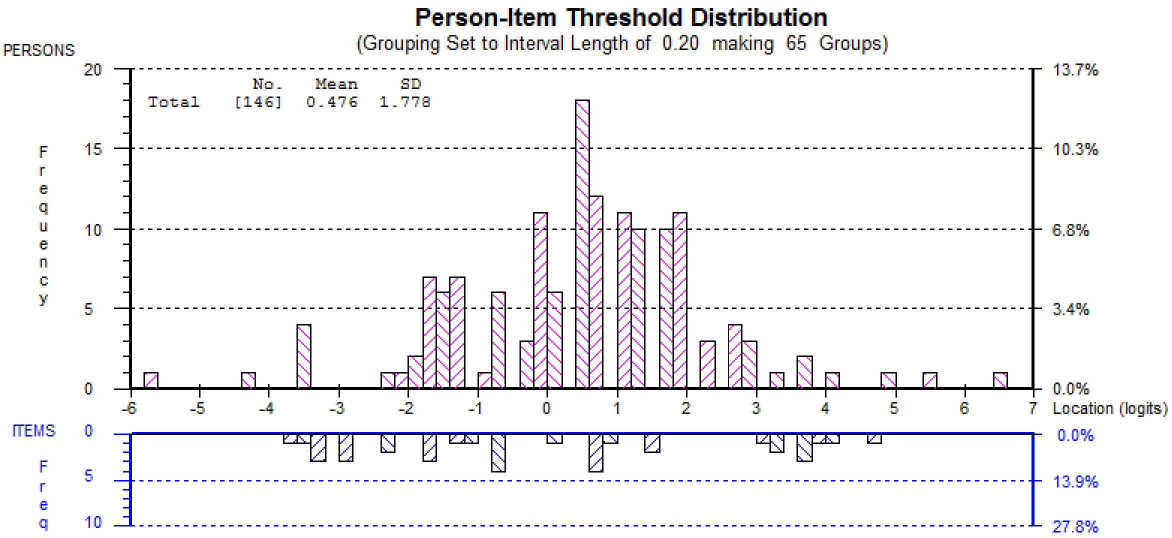


Figure 9-16 Person-Item threshold distributions for Scale 2.2, Model 2.2.2

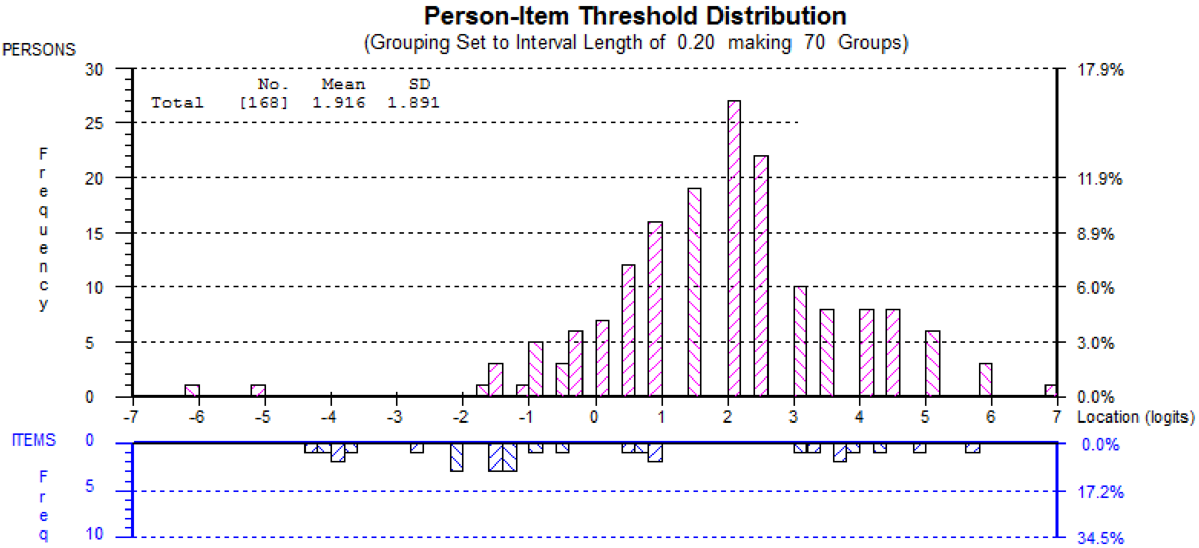
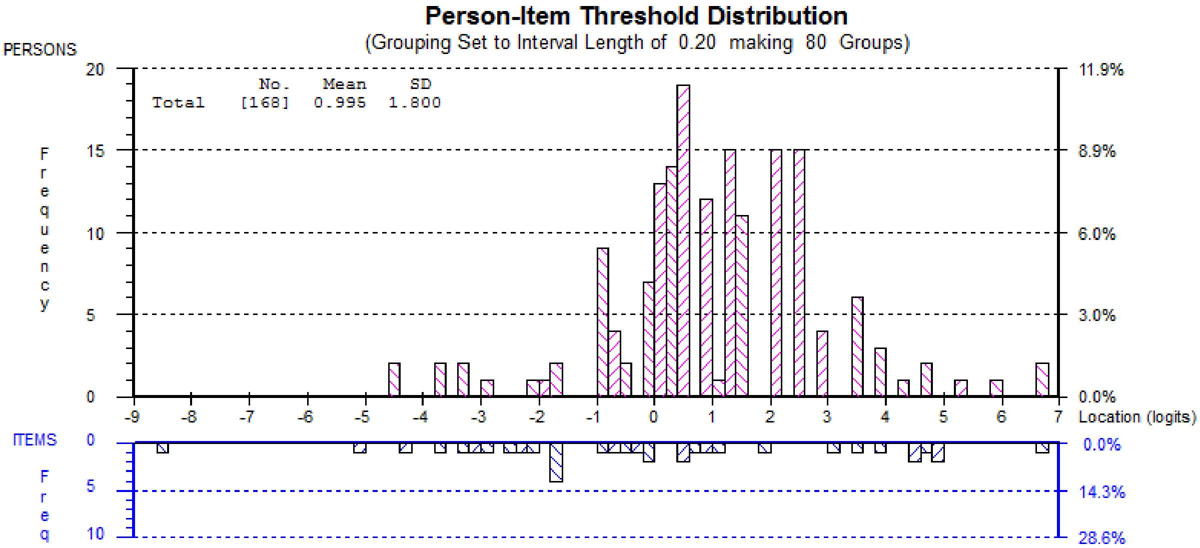


Figure 9-17 Person-Item threshold distributions for Scale 2.3, Model 2.3.2



9.6.3 Scoring (eHIQ-Part 2)

Person location values were transformed to raw scores using RUMM2030 software. As with the eHIQ-Part 1 sub-scales, raw scores were plotted against location values indicating raw scores are not linear across the full range of each scale. The eHIQ-Part 2 sub-scales however did show a more gradual increase in the slope across the full range in scores indicating more overall precision of measurement when compared to the eHIQ-Part 1 sub-scale scores (see Figure 9-18, Figure 9-19 and Figure 9-20). See Appendix 9.3 for raw scales transformation tables.

Figure 9-18 Scale 2.1 Location value from RUMM2030 against raw score

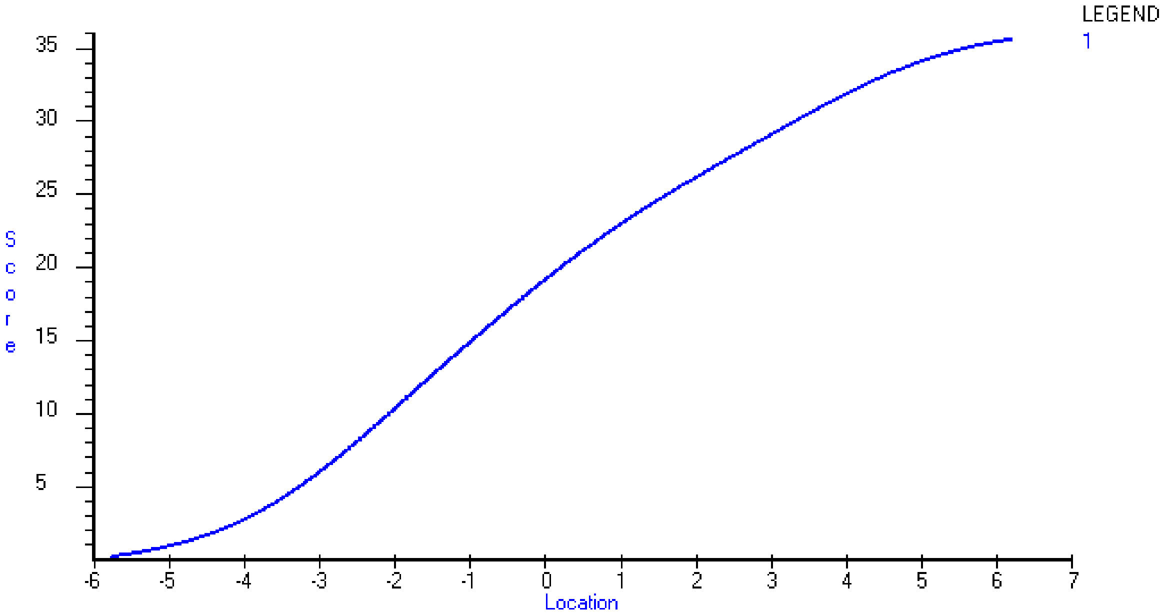


Figure 9-19 Scale 2.2 Location value from RUMM2030 against raw score

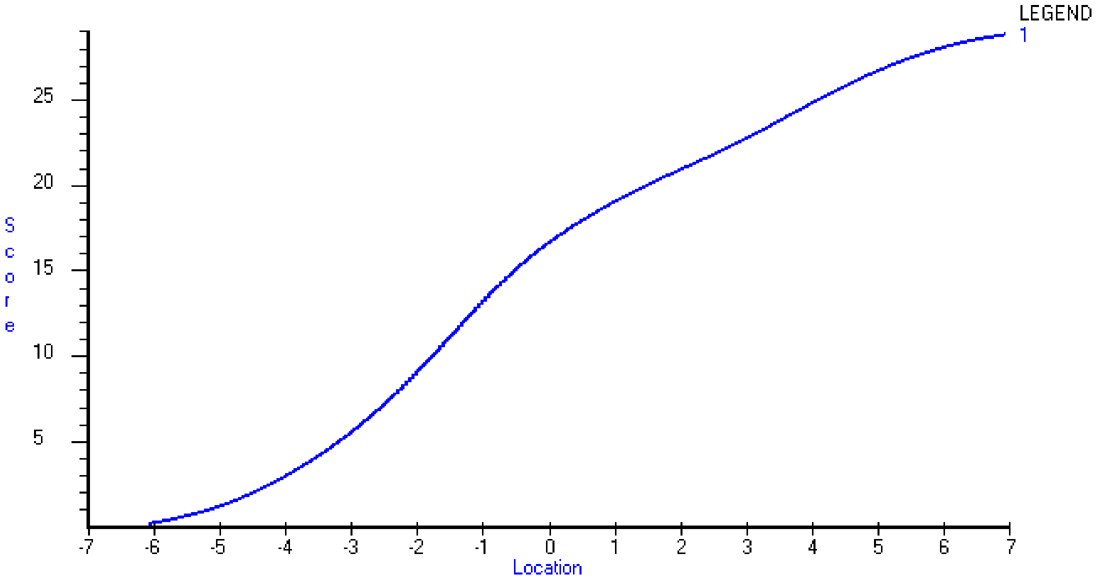
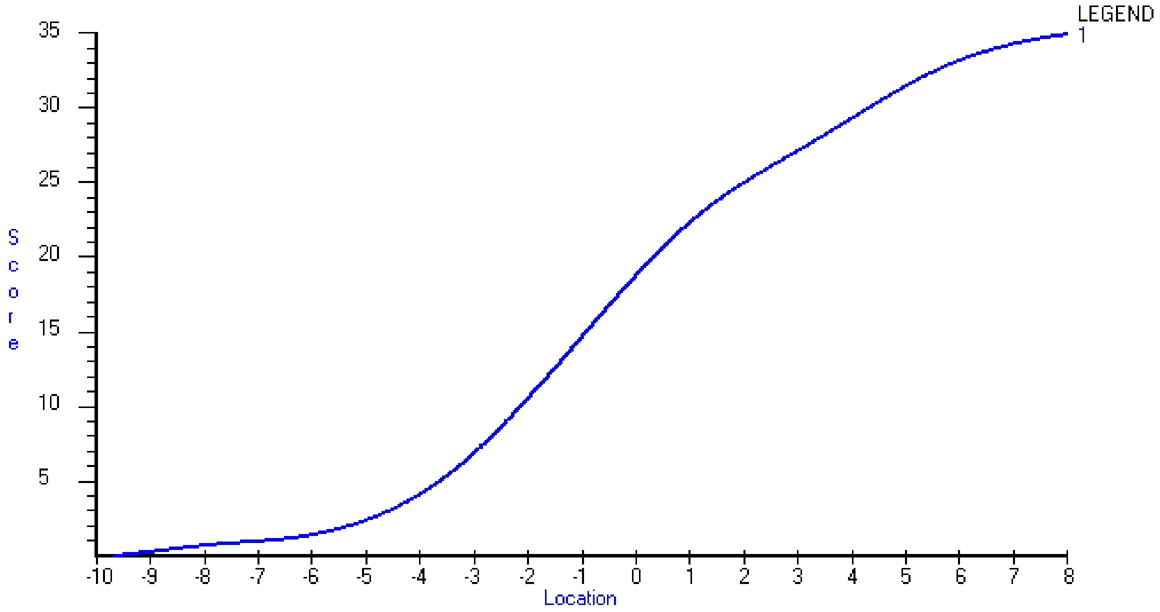


Figure 9-20 Scale 2.3 Location value from RUMM2030 against raw score



9.6.4 Rasch based scores versus classical test theory scores

Relationships between the eHIQ sub-scales and a range of population characteristics were investigated in Chapter 7 using CTT based scoring. These analyses were carried out again using Rasch based person location scores.

Parametric (t-test) and non-parametric tests were used to compare scores according to sex. As with the CTT scores, no significant differences were found between sex and all sub-scale scores using Rasch based person location scores. Parametric tests (Pearson correlation) also showed no significant differences between age and all sub-scale scores. Contrary to findings with CTT scores, no significant differences were found using non-parametric tests (Spearman Correlation) for age and 2.1 *Confidence and identification* (p=0.081). This is interesting, however, the differences found with CTT scores were very weak and would need further investigating with a larger sample size.

Parametric tests (One way ANOVAs) were carried out to test for significant differences among Rasch based scores and method of recruitment. The significance for the Homogeneity of variance statistic F was greater than 0.05 in all calculations indicating it was valid to conduct ANOVAs. As with the previous stage, there were significant differences found between the three recruitment phases and the sub-scale 1.2 *Attitudes towards sharing health experiences online* [F = 3.89, p=0.023]. Post-hoc comparisons using Tukeys HSD indicated that there were significant differences between those in Phase 1 (Mean=1.72, SD=1.36) and Phase 3 (Mean=2.99, SD=1.76).

With the exception of non-parametric tests for Scale 2.1 and age, all statistical comparisons between groups therefore resulted in similar results when using CTT scores or Rasch based person location scores.

9.7 Discussion

The analyses reported here aimed to confirm each sub-scales fit to the Rasch model to ascertain the assumptions of fundamental measurement were met: numerical order, possibility of addition and specific objectivity (i.e. items are independent of persons and vice versa) (Tennant et al., 2004).

For each eHIQ sub-scale, data were largely consistent with the expectations of the unrestricted Rasch model. Rescoring was necessary for one item in the eHIQ-Part 1 and four items in the eHIQ-Part 2. Overall summary fit statistics were reasonably good for each sub-scale and had good reliability with high PSI values. Independent t-tests of PCA components also demonstrated that each sub-scale was unidimensional. The transformation of the ordinal nonlinear scores for each sub-scale can now be converted into an equal interval metric.

The targeting of each sub-scale demonstrates items are reasonably well spread across the continuum with gaps in some scales indicating potential for improvement. Scale 1.2 demonstrated ceiling effects and a mean of 2.26 (SD 1.87). This indicates the scale may not be particularly well targeted for this sample, however, it would be unsurprising if those recruited for this study had a higher than average inclination towards sharing online health experiences compared to the general population. Both sub-scales in the eHIQ-Part 1 had a number of gaps along the item continuum. Although reasonably fitting Rasch model expectations, these sub-scales could potentially be improved in order to increase measurement precision around these locations.

Similarly to the sub-scales in the eHIQ-Part 1, Scale 2.2 measurement precision could potentially be improved at the positive end of the construct. Overall, however, the three sub-scales in the eHIQ-Part 2 are well targeted showing minimal floor and ceiling effects. This suggests the range of difficulty is accounted for within the range of the existing items.

Items were free from DIF with the expectation of eHIQ-Part 1, Item 7 (I would use the internet if I needed help to make a decision about my health (for example, whether I should see a doctor, take medication or seek other types of treatment)) for 'Sex'. It is unclear why there was bias in responses to this. One possible explanation for bias is that women report ill health more readily than men, despite having the same level of 'illness'. A further explanation may be that women find it easier than men to talk about personal topics in face-to-face conversations while men typically avoid self-disclosure (Coates, 2013). This may also carry over into the 'online' world. Item 7 also needed to be rescored to obtain ordered thresholds, so an underlying problem with this item may come apparent in further testing.

DIF was tested across a range of conditions and results suggested participants did not respond differently to items according to their health group. Sample sizes per condition however were small and future studies would need to obtain a larger sample size spread over a range of conditions to confirm this conclusion.

This chapter explored the suitability of using RMT to analyse the eHIQ scales. From the onset, the development of the eHIQ scales were developed using a CTT perspective. These Rasch analyses are encouraging as they demonstrate reasonable fit to the model, however, a perfect fit to the model was not found. RMT originated in the assessment of educational ability and has had success in the disability and health field since the late 1980's (Tennant et al., 2004). It is possible that data originating in these fields may be more amenable to satisfying the requirements of the Rasch model when compared to instruments assessing attitudes. For example, when assessing mathematically ability, it is perhaps more conceivable that examinees may have higher stakes in test outcomes and therefore may pay more attention or take more time over their answers (Curtis, 2004). Equally, it is possible that less attention may be paid when completing attitudinal scales as 'incorrect' answers have fewer consequences. Attitudinal scales may therefore be expected to have a higher proportion of spurious misftting persons (Curtis, 2004).

Unidimensional attitudinal scaling can also be difficult to achieve as attitudes by their very nature are often complex. Various factors may increase the magnitude of agreement or disagreement with each item, while the strength of agreement with other items may appear to be reduced due to issues such as social desirability (Curtis, 2004). Nonetheless the eHIQ appears to fulfil most criteria set by both CTT and RMT. This evidence suggests it is an instrument with robust psychometric properties.

It is worth noting that Rasch analysis adopts a different perspective to that of CTT as it is predicated on the measurement of ability (Rasch, 1960). It's origins lie primarily in educational testing, and hence ability is assumed to vary between individuals. Items are therefore endorsed, or not, dependent on 'item difficulty'. This measurement paradigm, differs fundamentally from that in CTT, subsequently became popular in disability research (Tesio, 2003, Belvedere and de Morton, 2010, Heinemann et al., 1993, Silverstein et al., 1991) and is now widespread in health and medical care (McHorney and Monahan, 2004, Tennant et al., 2004). CTT is primarily based on the assumption that attitudes can be measured (Thurstone, 1928), although it has also found application in the measurement of 'traits' such as personality (Cattell, 1943). This approach is also widely used in health and social sciences, and underpins the development of an enormous array of questionnaires measuring attributes as diverse as social function to physical ability. Whilst some authors have suggested results from one or other method may be more (or less) robust, what is often overlooked is their historical origins. The fact that a measure designed using CTT may not fit the Rasch model does not necessarily imply that the measure is of poor quality, but that it may be more appropriate to one approach than another. That said, data gained from CTT based on 'Likert scales' may be used as a basis for obtaining interval level estimates on a continuum by applying the Rasch model. Thus item groups may be created using CTT and tested using Rasch analysis, but they may not fit the model. It is interesting to note in this thesis that CTT seemed to provide more meaningful information than Rasch, quite possibly because the eHIQ measures phenomena that are typically closer to attitudes than ability per se.

The results of this chapter suggest that the eHIQ has items that can be successful analysed using both CTT and, for the most part, Rasch analysis. However, the CTT approach may be the most parsimonious given that the scale attributes were invariably good and provides for an easy to apply scoring metric. The next chapter will provide further analysis of the eHIQ comparing these two approaches.

CHAPTER 10: Comparing prototype websites using the eHealth Impact Questionnaire

10.1 Introduction

The previous chapters have documented the empirical research undertaken to develop the eHIQ using both traditional and modern psychometric methods. Sub-scales within the two-part questionnaire have been identified and their validity and reliability tested. In addition to individual sub-scale scores, a summary index score for each questionnaire part has also been established. Finally, the possibility of achieving equal interval scaling for each eHIQ sub-scale has been explored and raw scores can now be converted to locate an individual's location on each respective continuum. While it is important to establish the various sub-scales and possible scoring method, it is also important to explore the instruments performance in its intended research setting. This chapter documents the secondary analysis of a feasibility and acceptability focused randomised controlled trial (RCT) incorporating the use of the eHIQ. Data for secondary analysis was gathered for an exploratory trial conducted under the NIHR funded programme grant RP-PG-0608-10147 (EXPERT).

The overall aim of the RCT was to establish the feasibility of conducting a large scale online trial to examine the impact of health-related websites containing experiential information. Three exemplar conditions were selected to take part in the RCT and participants were allocated to either an active website containing experiential information or to a comparator health-related website containing scientific (or 'factual') information only. All design and multimedia features were similar with the only difference between the websites being the presence of experiential information on the active website. Data collection for one of the three exemplar conditions (people with asthma) had been completed at the time of this analysis and all analysis reported in this chapter therefore relates to this sub-sample. Analysis was also focused upon the performance of the eHIQ.

10.2 Objective

The aim of this chapter was to explore the performance of the eHIQ in a RCT setting. The objectives are outlined below:

- To explore data quality in terms of data completeness, score reliability and floor and ceiling effects.
- To provide preliminary descriptive statistics of participants completing the eHIQ.
- To compare eHIQ scores for the active group and the controlled group.
- To compare eHIQ CTT based scoring to eHIQ Rasch based scoring.

10.3 RCT design and methodology

All aspects of the RCT were carried out by a dedicated trial team based in the Nuffield Department of Primary Care Health Sciences, University of Oxford. Registered under the title *Experience of a Health Website Evaluated in a Randomised Trial* (Trial Registration Number: ISRCTN29549695), the 'EXPERT' trial followed a randomised controlled single-blind design. Three exemplar groups were chosen for recruitment and included asthma, smoking cessation and carers of individuals with multiple sclerosis. For each condition, two prototype websites were developed. The active website contained NHS Choices information plus patient experience information and the controlled website contained

NHS Choices health information, but did not provide experiential information. For full protocol see (Powell et al., 2012).

As described in the introduction, analysis in this chapter relates to a sub-sample (people with asthma) as data collection has been completed. While the primary objective of the trial was to explore feasibility issues in conducting this online trial design, secondary objectives included testing the administration of a range of self-reported measures to assess two styles of information, to pilot the eHIQ in an online trial and to explore potential effects on the various sub-groups included.

10.3.1 Participants

Males and females who were over 18 and living in England were recruited to the study. All participants were required to have internet access and experience of using the web. Patients who had been clinically diagnosed with asthma (as coded by their primary care electronic record) and prescribed inhaled corticosteroids for at least three months in the last year were invited to take part. All participants had to be willing and able to give informed consent to take part in the study.

Those excluded were those who were terminally ill or those who could not read English. Those with co-morbidities and those who were considered to be at risk if entered into the study by their GP were also excluded.

10.3.2 Recruitment

Participant recruitment was carried out through primary care. GP practices were provided with information about the study and asked to identify patients according to the criteria listed above. Eligible patients in participating GP practices were mailed a research invitation pack to be returned to the research team. Research invitation packs included a covering letter, patient information sheet, consent form, contact details form and reply-paid envelope. Upon receiving the consent form and contact details, participants were emailed a unique code with a link to the trial registration page.

10.3.3 Randomisation

Participants were asked to complete a set of baseline measures before they were randomised to either the intervention website or the comparator website. Following the electronic submission of the baseline measures, participants were randomised into the trial and allocated one of the two websites. Participants were randomised using a 1:1 ratio between the active and comparator websites. Participants created a unique password which enabled them to have unlimited access to the website they were allocated to over a period of two weeks. After two weeks participants were asked to complete follow up measures and access to the allocated website was withdrawn.

The trial was single blind as the trial team were blinded to which website each participant would be allocated during recruitment, however, due to the nature of the intervention and comparator website, participants could not be blinded. Nonetheless, participants were advised the study would be investigating how useful the health-related websites were and were not made aware of the trials specific objective of assessing experiential information.

10.3.4 Measure

Upon accessing the study registration page, participants were asked to complete several measures, one of which was the eHIQ-Part 1. After a period of two weeks, participants received an email asking them to complete follow up measures, including the eHIQ-Part 2.

10.3.5 Sample size

As the trial was exploratory in nature, the total planned sample size was 300 participants with equal numbers across the three population groups. Therefore, 100 participants were expected for the asthma sub-sample group.

10.3.6 Data analysis

Descriptive characteristics were used to describe the sample and report mean differences between the active intervention group and the comparator groups. Preliminary checks for each eHIQ sub-scale were carried out to ensure suitable scaling properties. Checks included interval consistency of the scales (Cronbach’s alpha >0.7), frequencies of missing data and floor and ceiling effects.

Comparisons between groups among eHIQ scores were examined using t-tests or ANOVAs. Continuous variables were examined using Pearson’s Correlation Coefficient. All analysis was carried out with both CTT and Rasch based scoring.

10.4 Results

Of the 200 people who had asthma who received a research invitation, 150 people completed the consent form and baseline questionnaires. Two participants withdrew prior to randomisation leaving 148 participants in total.

10.4.1 Characteristics and scale properties

A total of 148 people with asthma completed all or part of the eHIQ during the timeframe of the study. This sample size therefore exceeded the expected number of trial participants for the asthma sub-group. Of the 148 participants, 121 (81.8%) completed the eHIQ-Part 2 at the two week follow up. The mean age of the sample was 56.89 (SD 12.88, range 19-84 years). As was the case in the other studies reported in this thesis, more women (n=87, 58.8%) took part than men (n=61, 41.2%). Following randomisation, 75 people were allocated to the control website (health information only) and 73 people were allocated to the active website (health information plus experiential information).

Checks on each sub-scale showed good internal consistency (Cronbach's alpha ≥0.75) with all values exceeding the recommended value of 0.7 (see Table 10-1). Missing data was analysed to see if any item was causing particular difficulty for participants within the questionnaire. Of the 148 participants who completed eHIQ-Part 1, three people did not answer the following item:

- The internet is a good way of finding other people who are facing health-related decisions I may also face.

Of the 121 participants who completed the eHIQ-Part 2, one missing value was found for each of the following items:

- I value the advice given on this website.
- This website gives me confidence that I am able to manage my health.
- I feel I have a lot in common with other people using this website.

Table 10-1 **Cronbach's alpha values for eHIQ sub-scales**

Sub-scale	**n**	**Cronbrach's alpha**
1.1 Attitudes towards online health information	148	0.75
1.2 Attitudes towards sharing health experiences online	145	0.88
2.1 Confidence and identification	118	0.89
2.2 Information and presentation	121	0.78
2.3 Understanding and motivation	121	0.87

Each participant's eHIQ score was transformed into a score derived using the CTT based algorithim (0-100 metric) and, using a score derived from the Rasch raw score transformation to a person location estimate. Raw Rasch scores were converted to logit scores using the person location transformation tables in Appendices 9.2 and 9.3. Descriptive statistics demonstrated that all scales were reasonably normally distributed with some skewedness towards the higher end of the scales (see Table 10-2 and Table 10-3). No floor or ceiling effects were evident using both scoring systems.

Table 10-2 **eHIQ sub-scales and summary index scores (scores based on CCT)**

Sub-scale	n	Minimum	Maximum	Mean	SD	Skewness	Floor %	Ceiling %
1.1 Attitudes towards online health information	148	5.0	100.0	59.70	16.22	-0.28	0	0.7
1.2 Attitudes towards sharing health experiences online	145	4.17	100.0	62.01	17.93	-0.33	0	2.7
2.1 Confidence and identification	118	2.78	97.22	59.25	15.79	-0.49	0	0
2.2 Information and presentation	121	43.75	100.0	75.31	10.95	-0.27	0	1.7
2.3 Understanding and motivation	121	11.11	100.0	64.14	14.38	-0.69	0	1.7
eHIQ-Part 1 Summary	145	24.58	97.5	14.51	14.51	-0.12	0	0
eHIQ-Part 2 Summary	118	23.38	98.03	66.56	11.84	-0.40	0	0

Table 10-3 **eHIQ sub-scales (scores based on Rasch person location estimates)**

Sub-scale	n	Minimum	Maximum	Mean	SD	Skewness	Floor %	Ceiling %
1.1 Attitudes towards online health information	148	-4.56	3.36	-0.05	1.09	-0.004	0	0.7
1.2 Attitudes towards sharing health experiences online	145	-8.39	6.08	1.49	1.76	-0.71	0	2.7
2.1 Confidence and identification	118	-4.94	5.58	0.69	1.61	0.16	0	0
2.2 Information and presentation	121	-4.06	7.81	1.26	1.73	0.57	0	1.7
2.3 Understanding and motivation	121	-1.30	6.91	2.21	1.68	0.33	0	1.7

10.4.2 Comparisons between prototype websites, gender and age

10.4.2.1 Website allocation

Parametric (t-test) and non-parametric tests (Mann-Whitney U test of Significance) were carried out to investigate if there were significant differences according to website allocation and eHIQ scores. Mean differences between groups were small, but the control group had a less positive view of online health information in general (eHIQ-Part 1) and had a slightly more positive attitude (eHIQ-Part 2) towards the control website compared to those exposed to the active website. These observations were also supported using the Rasch based scoring. No significant differences were found between website allocation and eHIQ sub-scales scored using either CTT or Racsh based methods (see Table 10-4 and Table 10-5).

10.4.2.2 Gender

Parametric (t-test) and non-parametric tests (Mann-Whitney U test of Significance) were carried out to investigate if there were significant differences according to gender and eHIQ scores. Again, mean differences between groups were small, however, women were more positive towards online health information in general and towards the website they were exposed to during the intervention. These observations were consistent when using Rasch based scoring. No significant differences however were found between sex and all sub-scale scores regardless of the scoring system used (see Table 10-6 and Table 10-7).

10.4.2.3Age

Parametric (Pearson correlation) and non-parametric (Spearman Correlation) tests were carried out to investigate if there were significant differences in eHIQ score according to age. Weak negative relationships were found between attitudes towards online health information and age. Sub-scale 1.2 *Attitudes towards sharing health experiences online* and age had a significant negative relationship with age (n= -0.36, p<0.001) meaning that younger people had a more positive attitude towards using the internet for health information. A similar relationship was found using Rasch based scoring (r= -0.34, p<0.001). For all eHIQ-Part 2 sub-scales no relationship was found between attitudes to the specific website they viewed and the eHIQ sub-scales (see Table 10-8 and Table 10-9).

Table 10-4 Differences between allocation group and eHIQ CTT based score

Sub-scale	Website	n	Mean (SD)	Parametric		Non-parametric*
				t	p	p
1.1 Attitudes towards online health information	Control	75	59.27 (15.65)	-0.33	0.745	0.684
	Active	73	60.14 (16.87)			
1.2 Attitudes towards sharing health experiences online	Control	72	61.86 (16.85)	-0.10	0.922	0.724
	Active	73	62.16 (62.16)			
2.1 Confidence and identification	Control	57	59.80 (17.03)	0.36	0.719	0.493
	Active	61	58.74 (14.67)			
2.2 Information and presentation	Control	59	76.06 (11.38)	0.73	0.465	0.477
	Active	62	74.60 (10.56)			
2.3 Understanding and motivation	Control	59	64.31 (15.22)	0.13	0.899	0.741
	Active	62	63.98 (13.66)			
eHIQ-Part 1 Summary	Control	72	61.07 (13.55)	-0.03	0.975	0.672
	Active	73	61.15 (15.48)			
eHIQ-Part 2 Summary	Control	57	67.33 (12.46)	0.67	0.501	0.379
	Active	61	65.85 (11.29)			

*Non-parametric test: Mann-Whitney U test of Significance

Table 10-5 Differences between allocation group and eHIQ Rasch based score

Sub-scale	Website	n	Mean (SD)	Parametric		Non-parametric*
				t	p	p
1.1 Attitudes towards online health information	Control	75	-0.10 (1.09)	-0.63	0.533	0.623
	Active	73	0.01 (1.10)			
1.2 Attitudes towards sharing health experiences online	Control	72	1.49 (1.52)	-0.00	0.998	0.724
	Active	73	1.49 (1.99)			
2.1 Confidence and identification	Control	57	0.75 (1.75)	0.42	0.679	0.493
	Active	61	0.63 (1.48)			
2.2 Information and presentation	Control	59	2.36 (1.78)	0.96	0.341	0.755
	Active	62	2.06 (1.58)			
2.3 Understanding and motivation	Control	59	1.29 (1.80)	0.21	0.836	0.425
	Active	62	1.22 (1.67)			

*Non-parametric test: Mann-Whitney U test of Significance

Table 10-6 **Differences between sex and eHIQ CTT based score**

Sub-scale	Sex	n	Mean (SD)	Parametric		Non-parametric*
				t	p	p
1.1 Attitudes towards online health information	Male	61	58.69 (16.28)	-0.63	0.529	0.763
	Female	87	60.40 (16.23)			
1.2 Attitudes towards sharing health experiences online	Male	59	60.45 (17.45)	-0.87	0.388	0.316
	Female	86	63.08 (18.29)			
2.1 Confidence and identification	Male	43	58.91 (18.20)	-0.18	0.862	0.902
	Female	75	59.44 (14.37)			
2.2 Information and presentation	Male	45	73.82 (11.91)	-1.15	0.251	0.108
	Female	76	76.19 (10.32)			
2.3 Understanding and motivation	Male	45	63.15 (16.63)	-0.58	0.561	0.681
	Female	76	64.73 (12.96)			
eHIQ-Part 1 Summary	Male	59	60.14 (13.47)	-0.66	0.508	0.510
	Female	86	61.77 (15.22)			
eHIQ-Part 2 Summary	Male	43	66.04 (13.20)	-0.36	0.178	0.509
	Female	75	66.86 (11.06)			

*Non-parametric test: Mann-Whitney U test of Significance

Table 10-7 Differences between sex and eHIQ Rasch based score

Sub-scale	Sex	n	Mean (SD)	Parametric		Non-parametric*
				t	p	p
1.1 Attitudes towards online health information	Male	61	-0.17 (1.10)	-1.13	0.261	0.554
	Female	87	0.03 (1.08)			
1.2 Attitudes towards sharing health experiences online	Male	59	1.39 (1.51)	0.55	0.566	0.316
	Female	86	1.56 (1.92)			
2.1 Confidence and identification	Male	43	0.68 (1.89)	-0.05	0.959	0.902
	Female	75	0.70 (1.43)			
2.2 Information and presentation	Male	45	2.00 (1.87)	-1.06	0.294	0.669
	Female	76	2.33 (1.55)			
2.3 Understanding and motivation	Male	45	1.18 (1.99)	-0.37	0.710	0.106
	Female	76	1.30 (1.56)			

*Non-parametric test: Mann-Whitney U test of Significance

Table 10-8 **Relationship between age and eHIQ CTT based score**

Sub-scale	n	Parametric		Non-parametric	
		Pearson r	p	Spearman Correlation	p
1.1 Attitudes towards online health information	148	-0.13	0.105	-0.16	0.058
1.2 Attitudes towards sharing health experiences online	145	-0.36	0.001*	-0.34	0.001*
2.1 Confidence and identification	118	-0.01	0.928	0.03	0.767
2.2 Information and presentation	121	-0.09	0.344	-0.16	0.085
2.3 Understanding and motivation	121	-0.00	0.962	0.01	0.939
eHIQ-Part 1 Summary	145	-0.32	0.001*	-0.28	0.001*
eHIQ-Part 2 Summary	118	-0.04	0.665	-0.04	0.661

*Significant at <0.05 level (2 tailed)

Table 10-9 **Relationship between age and eHIQ Rasch based score**

Sub-scale	n	Parametric		Non-parametric	
		Pearson r	p	Spearman Correlation	p
1.1 Attitudes towards online health information	148	-0.16	0.053	-0.16	0.053
1.2 Attitudes towards sharing health experiences online	145	-0.34	0.001*	-0.339	0.001*
2.1 Confidence and identification	118	0.03	0.767	0.028	0.767
2.2 Information and presentation	121	-0.15	0.098	-0.004	0.964
2.3 Understanding and motivation	121	-0.00	0.964	-0.151	0.098

*Significant at <0.05 level (2 tailed)

10.5 Discussion

This chapter aimed to document the pilot testing of the eHIQ in an applied research setting. The analysis sought to provide preliminary findings of scaling properties and to examine the questionnaires ability to pick up differences between population groups. CTT and Rasch based scoring were used in all analyses.

Analysis highlighted good scaling properties in relation to internal consistency (Cronbach's alpha ≥0.75) and no floor or ceiling effects. Extremely low numbers of missing data were observed. Items with missing data appeared to be largely 'missing at random' suggesting this there were no problems with particular items (Rubin, 1976). These observations suggest participants find completing the questionnaire reasonably easy and the range of measurement is adequate for the population group. Items also appear to be measuring the same unidimensional construct in each sub-scale as supported by the good levels of internal consistency.

Both CTT based and Rasch based scoring algorithms were adopted. Results indicated both scoring systems performed very similarly. The eHIQ scores did not show significant differences between the control and active websites. This RCT was carried out to assess the feasibility of conducting a larger, sufficiently powered trial. Therefore it was unlikely to have shown a significant difference in scores. It is also unsurprising no significant differences between groups were found as a relatively short period of time was given to participants to use the websites. Furthermore, analysis was only carried out on one of the three exemplar groups as data collection was incomplete for the remaining two exemplar groups (Carers of people with MS and those who wished to stop smoking). Further analysis may be carried out when all data is collected for the RCT.

Some promising results were demonstrated in relation to the questionnaires sensitivity. Although a statistically significant difference was not found, women were more favourable towards general and experiential online health information (eHIQ-Part 1) and towards the specific website they were allocated (eHIQ-Part 2). Younger people were also inclined to have a more positive attitude towards using the internet in relation to health. This was found to be significant in the case of sharing experiential information online. These findings support construct validity given existing literature which suggests more women than men are likely to use the internet in relation to health and younger people are more open to using the internet in relation to health compared to older people (Fox and Duggan, 2013).

Overall, when comparing CTT based scored compared to Rasch based scores, this study did not highlight many benefits from using one form of scoring over the other. The overall recommendation would therefore be to use the CTT based scores for future research as scores may be considered easier to calculate.

CHAPTER 11: Discussion and conclusions

11.1 Introduction

The healthcare landscape is transforming with the increasing use of the internet (Reis et al., 2013). Health-related websites have developed to be much more than information sites: they are used to exchange experiences and find support as well as information and advice. Research to date has typically focused on the quality and trustworthiness of online health information, or compared online information sources with more traditional information sources (such as, leaflets or pamphlets). While these issues are important, they cannot identify how the experience of using particular websites and their content may *affect* the user. Some evidence has shown that web-based interventions have positive outcomes (for example, increased patient understanding and empowerment), however further research is needed to fully understand the implications (Samoocha et al., 2010).

It is essential that research is conducted to inform health professionals and web developers how to maximise the positive effects of using various forms of online health information. To do this, a robust research design is required using a suitable instrument to compare alternative forms of presenting health information. Following preliminary literature searching, this thesis identified a need to develop an appropriate tool with which to measure the impact of using a health-related website.

11.2 The eHealth Impact Questionnaire

This thesis has documented the stages taken to develop the eHealth Impact Questionnaire (eHIQ). The methods used reflect best practice guidelines in health-related questionnaire development (FDA, 2009, EMA, 2004) and, where possible, adhere to design recommendations for web surveys (Dillman and Bowker, 2001, Dillman, 2011, Toepoel et al., 2009).

At the forefront of the development process was the premise that the instrument should be applicable for use among a range of health groups and be suitable for examining various styles of health-related websites. Therefore, during all stages of the study, the items and questionnaire design were tested for suitability among individuals viewing a website which was aimed at: 1) People with long term conditions, 2) People interested in health promotion or screening information and, 3) Carers of people with a health condition. The questionnaire was also tested on websites containing: 1) Standard 'scientific' health information, 2) Experiential health information and, 3) Forums and blogs.

11.2.1 Psychometric properties

11.2.1.1 Content and face validity

Several steps were taken to incorporate users' perspectives to promote good measurement properties. The item pool was primarily guided by the theoretical findings of a recent literature review (Ziebland and Wyke, 2012). The review drew on existing literature to identify the potential positive and negative effects of seeking and sharing online experiential information across a range of health groups. As such, it was considered highly relevant and informative to the early stages of the questionnaire development. The item pool was also informed by a secondary analysis of interview transcripts from the

Health Experiences Research Group (HERG) archive. Access to these existing interviews relating to carer and patient experiences of health provided a rich and detailed data source with which to conduct a focused in-depth analysis into the impact of using health-related websites. Analysis also enabled the incorporation of internet users' actual words as a basis for questionnaire items enhancing face and content validity.

Accounts of using the internet in relation to health were evident in 93 transcripts (which included 99 participants) of the 198 HERG transcripts screened. Men and women reported accessing health-related websites intermittently with frequency of use peaking according to key health events (such as diagnosis, or progression of an illness). Participants described using a range of online resources (including conventional health websites, health discussion forums and blogs) and many combined the information they found online with advice from healthcare professionals. This was in keeping with previous research which has indicated that people value various forms of online information and use the web in addition to seeking advice from healthcare professionals, not as a substitute (McMullan, 2006, Fox, 2011b, Wyke et al., 2011, Entwistle et al., 2011).

Analyses demonstrated that preferences for different types of online health information varied between participants. This was considered an individual's 'baseline' feeling towards using the web for health information and was hypothesised to have an effect on the impact of using a specific website. Research has indicated that attitudes towards using the internet are significantly associated with intention to use the internet for health information (Kim et al., 2012). These findings were largely based upon modified versions of instruments which were not originally developed for use in relation to health (Davis, 1989, Compeau and Higgins, 1995) and, with the exception of one of the instruments used, *The Internet (iKnow) measure* (Potosky, 2007), were not developed in the more recent Web 2.0 climate. Therefore, it was felt that developing a scale to assess general attitudes towards online health information would be useful in future research (the eHIQ-Part 1).

To inform the second pool of items relating to using a specific health-related website (eHIQ-Part 2), themes from the secondary data analysis were mapped onto the domains identified in Ziebland and Wyke's (2012) literature review. The literature review identified seven domains relating to the potential effects of using online experiential information. The secondary analysis supported these findings. Five domains were mapped onto themes relevant to the impact of using various styles of health-related websites across a range of conditions, whilst the remaining two domains were embedded across these five domains.

The first theme, *Information*, incorporated learning about a health condition, being informed of what to expect and recognising decisions which may need to be made in the future. Health-related websites were also a valuable resource for finding practical tips and advice to help manage health. Negative consequences included feeling overwhelmed with the amount of information available, feeling uncertain of what information to trust and finding upsetting information.

The second theme, *Feeling supported,* related to finding support through others going through similar health experiences. In some cases, accessing a particular health-related website was the first time participants learnt that other people had similar health concerns. Increased hope and confidence was also found through gaining and lending support. Some participants however did not find it useful to contrast their health experiences with others and occasionally found sensitive issues upsetting.

The third theme, *Relationships with others,* referred to feeling less isolated and that others understood what they were experiencing (for example, in an online community). Using websites also had the potential to ease interpersonal relationships as it provided an area where users could discuss their health concerns with people going through similar experiences and learn to articulate how they were feeling to relatives and friends. Online support however can occasionally lead to alienation as one respondent indicated they felt hurt that their partner found it easier to discuss her feelings with others online.

The fourth theme, *Experiencing health services,* concerned the use of websites when deciding whether medical advice should be sought. Use of the internet could help participants navigate the health system and helped to prepare people for upcoming decisions. Participants sometimes discussed using the internet to corroborate advice, to find out more information and to reflect on decisions made after accessing health services. In addition, the Northumbrian focus group transcripts highlighted some concerns over the misinterpretation of information which may result in discouraging patients with legitimate health needs approaching healthcare professionals.

The final theme, *Affecting behaviour,* incorporated the motivational effect some website content may have on the individual. For example, learning of the consequences of ignoring best practice when managing a long term condition encouraged people to improve their own healthcare regime. In the User Panel discussion group however, some concern surfaced regarding the potential for some people to be influenced to take part in unhealthy behaviours (for example, when using pro-anorexia websites).

The five outlined themes were consistent with the domains published in the literature review providing a strong basis for the conceptual underpinning of the eHIQ-Part 2 item pool. Issues relevant to the remaining two domains in the literature review were present throughout the five confirmed themes. For example, *Learning to tell your story* was incorporated in *Relationships with others* when a participant discussed being able to communicate more succinctly with their partner after using an online forum or talk to doctors. *Visualising Disease* was also incorporated throughout the five themes identified when participants described images reinforcing people’s stories or messages and their potential negative impact when seeing images of people with a degenerative condition. It was therefore felt many of the issues addressed in the two additional domains were adequately represented.

Secondary data analysis of the HERG transcripts proved to be an invaluable way of informing the item pool. When constructing an item pool, it is common to gain a detailed understanding of a construct through conducting focused interviews with the relevant population. This questionnaire however needed to be applicable across a range of conditions and roles (i.e. people with a health condition and carers). Conducting in-depth interviews to explore the effects and experiences of using online health information across many different health groups would have been unrealistic in the timeframe available. Using the available HERG archives was therefore a crucial component in ensuring the users perspective was captured in the item pool. Some limitations surrounding secondary data analysis should however be acknowledged.

A potential limitation of analysing existing qualitative datasets is that the analyst cannot follow up relevant points raised by the participant or explore emerging themes in subsequent interviews (Pope et al., 2000, Heaton, 1998, Ziebland and Hunt, 2014). Obtaining contextual knowledge surrounding the interviews and becoming familar with a large dataset can also be challenging and time consuming when selecting the required sample. In this research, such challenges were somewhat alleviated with the availability of interview video recordings of the original interview and through liaising with the primary interviewers. Furthermore, cross checking and following up identified themes with the Northumbrian focus group transcripts and the user panel forms provided strong evidence for content validity and assurance that no new themes had been overlooked.

Expert and patient opinion further corroborated that the items covered themes adequately and ensured items were acceptable to patients and carers. Members of the expert panel comprised of clinicians and academics in the field of eHealth and agreed candidate items adequately covered the potential effects of using the internet for health information. Patient opinion through cognitive interviewing also allowed the opportunity to ask participants if they thought important issues had been left out. These steps supported content and face validity of the item pool and informed the appropriateness of the web survey design.

11.2.1.2 Item reduction, validity and reliability

Two independent item pools were confirmed as suitable to enter psychometric testing. The first item pool (eHIQ-Part 1) contained 23 items asking respondents about their general attitudes towards health-related websites whilst the second item pool (eHIQ-Part 2) contained 39 items asking the respondent about their attitudes towards a specific health-related website. The online questionnaire was administered in a pilot survey (Chapter 6) and items were removed using a series of a-priori item reduction decisions and exploratory factor analysis.

The reduced, 48 item questionnaire was subsequently administered in Stage Four (Chapter 7) to explore the validity and reliability of the new measure. Analyses resulted in the removal of three items from the eHIQ-Part 1 and confirmed the presence of two sub-scales: 1.1) *Attitudes towards online health information* and 1.2) *Attitudes towards sharing health experiences online*. Eight items were removed from the eHIQ-Part 2 and three sub-scales were retained: 2.1) *Confidence and identification*, 2.2) *Information and presentation* and 2.3) *Understanding and motivation*. The two sub-scales in the eHIQ-Part 1 explained 56.6% of the variance while the three sub-scales in the eHIQ-Part 2 explained 61.9% of the variance.

The sub-scales identified were considered conceptually relevant to the impact of using health-related websites. The first eHIQ-Part 2 sub-scale, *Confidence and identification*, measures the extent to which an individual identifies with others using a specific website and the extent to which the website makes them feel confident in managing and discussing their personal health with others. These issues were found to be important within the secondary analysis of HERG transcripts and are further supported through research relating to patient empowerment when using health-related websites, particularly in research relating to online support groups (van Uden-Kraan et al., 2008, Mo and Coulson, 2010).

The second sub-scale, *Information and presentation*, assesses the extent to which an individual feels a specific website is easy to use and contains a range of easy to understand information which appears trustworthy. The sub-scale also asks about the appropriateness of images used. Although they may appear to be slightly superficial elements, the decision to trust a specific website can be made very quickly and trust perceptions have been linked to engagement with websites in previous research (Sillence et al., 2007b, Rains, 2007). The elements measured by the *Information and presentation* scale are therefore known to be important issues which contribute to the user experience.

The final eHIQ-2 sub-scale, *Understanding and motivation*, measured the extent to which the respondent felt reassured, understood their condition better and felt motivated to manage their health after viewing a specific website. These components were reflected in the HERG transcripts and web-based interventions have also been shown to encourage positive behavioural outcomes through improved knowledge (for example, through increasing knowledge relating to nutrition or asthma treatment) (Wantland et al., 2004).

Statistical analyses confirmed the hypotheses that the eHIQ sub-scales were moderately related to the chosen reference measures of similar constructs. Correlations between the eHIQ-Part 1 and the *HINTS* trust item were weak to moderate ranging from -0.35 to -0.52 while correlations between the eHIQ-Part 2 and the modified *Access to quality information* scale were moderate to good ranging from 0.75 to 0.76. This meant the eHIQ sub-scales were sufficiently related to each respective reference measure to demonstrate convergent validity, yet sufficiently dissimilar to validate the independent construct (Hinkin, 1998). The scaling properties of the eHIQ sub-scales were also satisfactory and demonstrated little or no floor and ceiling effects. All sub-scales had some evidence of being negatively skewed indicating that the distribution was concentrated on the positive end of the scale (i.e. positive view of the website). This was expected as participants were asked to view established websites run by government organisations (for example, NHS Choices) or major charities (for example, the MS Society) and likely to encourage favourable results. Despite the concentration of scores at the positive end of the construct, the distributions were relatively normal. One consideration for the future development of the eHIQ may be to test the merits of increasing the number of positive response options to reduce potential bias (Streiner and Norman, 2008).

All sub-scales showed good internal consistency achieving a Cronbrach's alpha of ≥ 0.77. This gave a good indication that all items within each sub-scale were measuring the same latent variable. Good test-retest reliability was found among those who completed the eHIQ again after a two week interval (n=83). Intra-class correlation coefficients ranged from 0.76 to 0.91 for all sub-scales. This indicated that, providing there had been no change in the construct being measured, the items in each eHIQ sub-scale are answered consistently over time (Jenkinson and McGee, 1997, Hays and Revicki, 2005).

Higher order factor analyses confirmed the appropriateness of combining sub-scales to create summary index scores for each questionnaire part. Creating a summary index score can reduce the number of comparisons during statistical analyses and hence the likelihood of findings due to chance. A short form version of the questionnaire was also successfully created to reduce responder burden which may be particularly useful in studies which incorporate a large number of measures. High levels of agreement between each summary score and their respective short form score was also demonstrated (ICC ≥0.91) which is encouraging for those who may wish to use the short form questionnaire in their research.

A simple scoring algorithim was used to create sub-scale and summary index scores on a 0-100 metric. Scoring was further explored using Rasch Measurement Theory (Chapter 9). When data fits the Rasch model equal interval level measurement can be assumed (Wright and Stone, 1979, Andrich, 2011). Sub-scales were largely found to fit the Rasch model, although Scales 1.2 and 2.2 had slightly high mean person location scores and some ceiling effects. The analysis further indicated some gaps in measurement on the continuum for each construct which may compromise precision. This was particularly relevant for the lower end of Scale 1.2, *Attitudes towards sharing online health information* (Chapter 9, Table 9-6), and at a gap (between logit 1.0 and 3.0) in the positive end of Scale 2.2, *Information and presentation* (Chapter 9, Table 9-16).

Analyses in Chapter 10 used both Classical Test Theory based scores and Rasch based scoring to compare differences between the impact of using a website containing 'scientific' health information only (control website) and a website containing 'scientific' health information plus experiential information (active website). The opportunity to conduct secondary analysis on available data from a randomised controlled trial (RCT) was very useful to explore the eHIQ's performance in a setting for which it was designed. The RCT was carried out independently from the work documented in this thesis, however, was funded by the same NIHR Programme grant (RP-PG-0608-10147).

The RCT aimed to explore the feasibility and acceptability of assessing the effects of experiential information in health-related websites across three exemplar health groups. At the time of the secondary analysis, data collection had been completed for one of the exemplar health groups (people with asthma). Analyses of the available data demonstrated no significant differences between eHIQ sub-scale scores according to allocation group. This was unsurprising given the relatively short period of time given to participants to use the website and the sample size. High completion rates however supported the questionnaires acceptability in an RCT. Some promising results were also demonstrated in relation to the questionnaires sensitivity. For example, women favoured online information more than men (eHIQ-Part 1) and scored more positively than men in relation to the specific website viewed (eHIQ-Part 2). Younger people also had a more positive attitude towards using the internet in relation to health, and significantly so for sharing experiential information online. These results supported prior hypothesis based on what is known about the typical characteristics of those seeking online information (Fox and Duggan, 2013, Atkinson et al., 2009, Ybarra and Suman, 2006) and therefore support the face, content and construct validity of the questionnaire. The results give some indication that, given the correct sample size, the eHIQ will be sensitive to picking up differences between groups.

Determining what would be an adequate sample size to be confident of statistical inferences using the eHIQ may be explored in future research. One way in which this may be investigated would be ascertain what a meaningful change in eHIQ scores is likely to be. Participants may be asked to view what the literature would suggest is a ‘poor’ website (for example, the presence of negative trust cues such as pop-up advertisements) and then complete the eHIQ. This could be followed by asking participants to view what the literature would consider a ‘good’ website (for example, positive trust cues such as quality indictor markers) and then complete the eHIQ. The difference in the scores may be considered a meaningful difference and be used to carry out a sample size calculation.

Further testing of the eHIQ sub-scales may also be carried out when full data completion of the RCT described in Chapter 10 has been carried out. This may provide an opportunity to complete a Confirmatory Factor Analysis (CFA) on the sub-scales identified with the eHIQ. CFA can be used to assess the extent to which items conform to a-priori domains in the data set (Kline, 2000, Nunnally and Bernstein, 1994). The small sample size of the asthma sub-sample precluded the undertaking of a CFA (Kline, 2000) and, as the instrument is designed for use across conditions, CFA should ideally be carried out using data from a range of health groups.

Overall, the eHIQ sub-scales are highly relevant to attitudes towards using the internet for health information (eHIQ-Part 1) or using a specific health-related website (eHIQ-Part 2). Existing measures which have been used to assess the impact of websites to date have not been designed specifically for use in relation to the internet or only cover very specific aspects of using a website. For example, the instruments identified in Chapter 3 were originally designed for GP consultations (*Patient Enablement Instrument*) (Yardley et al., 2010) or aspects of trust (*Web Trust Questionnaire)* (Harris et al., 2011).

A further strength of the eHIQ is that it can be used across conditions and has incorporated the potential ‘user’ in all stages of development. This is in contrast to other instruments which try to include the views of the patient through the use of existing literature, but fail to directly include the patient in the development process, for example, *Information Comprehensiveness Tool (ICT)* (Warren, 2014). One further interesting instrument which has come to light during the period of this research is the *Information assessment Method for all (IAM4all)* (Pierre et al, 2014). The IAM4all is designed to evaluate outcomes associated with seeking and receiving online health information. This tool is limited in that it is theory driven and did not include potential users at the initial stages of development. Items were developed by researchers using literature relating to 1) Situational relevance (Acquisition), 2) Cognitive/Affective Impacts (Cognition), 3) Information use (Application) and 4) Health benefits. However, items were tested among a group of 16 lay participants and subsequently among an expert panel. The IAM4all appears to be useful in website evaluation studies and is yet to be validated using statistical analysis. In due course it may be interesting to compare the performance of the IAM4all and the eHIQ.

Developing generic health-related questionnaires is now common place in health research (Bowling, 2005). This research incorporated a range of conditions in each stage of development. Analysis demonstrated that developing an appropriate generic questionnaire for use with websites containing scientific 'facts and figures' information and experiential information was feasible. Further testing of the questionnaire however would be beneficial among groups which were particularly difficult to recruit at various stages of this research, for example, carers of people with a health condition. It is also worth noting that, as with other generic health-related questionnaires, the suitability of using the eHIQ in a specific study will depend on the study's research questions. For studies which aim to answer very specific questions which are particularly characteristic to the chosen population, a condition-specific instrument may be more appropriate for the research context.

11.2.2 Using the eHealth Impact Questionnaire

The eHIQ is designed for use in studies comparing various forms of health-related websites and it is recommended that the scoring system based upon Classical Test Theory is routinely used. The eHIQ-Part 1 may be used on a longitudinal basis, however, as items are tapping into underlying individual information preferences, it is unlikely scores will change significantly over a relatively short period of time. Use in a cross-sectional study is therefore more appropriate. The eHIQ-Part 1 may be useful for studies which require a baseline knowledge of the samples overall attitude towards using online health information. Preliminary findings in this thesis indicate the eHIQ-Part 1 does not have ceiling or floor effects, therefore, it may be useful to examine the spread of scores using the standard deviation.

The eHIQ-Part 2 is suitable for use after a participant has used a specific health-related website. The questionnaire should ideally be used in controlled comparison trials however, it may potentially be used for simple website evaluation purposes. The lack of missing data throughout the development stages and pilot RCT indicate that the eHIQ items are acceptable to participants and are likely to be applicable to several different health groups and carers. The pilot RCT outlined in Chapter 10 also demonstrated the measures satisfactory performance in its intended research setting. Such observations are encouraging for the use of the eHIQ in trials of a similar nature in the future.

11.2.3 Recruitment for online surveys

Throughout this research, women, younger people and those with a reasonably high level of education were the most common participants. This reflects the characteristics of those using the internet in relation to their health (Ybarra and Suman, 2006, Powell et al., 2011, Atkinson et al., 2009, Fox and Duggan, 2013). As such, the populations sampled were likely to be reflective of the most likely users. While these results were promising for the representativeness of the study populations, however, a number of difficulties are associated with recruitment strategies for online surveys. Best practices for implementing online surveys and calculating response rates are less established when compared to routine practices for administering paper based surveys. This thesis explored numerous issues which proved to be challenging when using the internet to conduct research (for example, the presentation of recruitment materials, methods of recruitment and achieving a 'representative sample').

Recruitment proved to be extremely difficult throughout the developmental stages of the eHIQ. Poor recruitment rates may have been experienced as no 'incentives' were used to encourage participation (for example, monetary incentives or the opportunity for exclusive access to a new website). Recruitment for the development of the eHIQ largely relied on the altruistic behaviour of potential participants. While this approach added legitimacy to the sample by providing less incentive for multiple questionnaire completion, it may have been less enticing for people to take part. The difficulties experienced with recruitment in the developmental stages were in contrast with the RCT recruitment which had a 75% response rate. This may, in part, have been due to participants having the incentive of being given access to a 'new' website. Other online studies which have provided access to new websites have also experienced high rates of recruitment (Wallace et al., 2011).

Recruitment was also difficult during the developmental stages as a number of health groups were considered ‘hard-to-reach’. Several charities advised many of their members were not frequent users of the internet (for example, carers connected to the MS Society) and some were quite difficult to contact (for example, people who wish to stop smoking were difficult to pinpoint using NHS Read codes).

A secondary aim of Stage Four (Chapter 7) was to explore paper based recruitment and online recruitment. Potential participants were asked to take part in the research through a mailed research invitation, through an online research invitation sent to their email inbox, or through an open posting of a research advertisement on a relevant website. Of those recruited in Phase 1 (mail out recruitment) and Phase 2 (online recruitment), the majority of participants were recruited online (n=95, 68.35%). It is possible to speculate that online recruitment was more fruitful because: 1) Online recruitment was more user friendly as participants simply clicked on an electronic link to access the study website, 2) All those recruited online inevitably had access to the internet while a proportion of those in the mail out recruitment group were likely to have no access to the internet.

Despite the attraction of online recruitment, it has inherent difficulties when considering the representativeness of the sample. It was impossible to target potential participants in each health group across a range of demographic characteristics and calculating an accurate response rate was extremely difficult. Estimates were made for those who received a research invitation through an email sent their inbox (MND, n=32, 12.8% and MS Society branches n=6, 4%). Online research advertisements were also placed across a range of online resources (for example, Facebook and websites belonging to national organisations) in an attempt to capture a range of characteristics within the sample. From the group recruited through online methods, 225 people viewed the front (title) page of the web survey. As 96 people completed the questionnaire, the ‘response rate’ was estimated to be 42.7%.

In contrast, recruiting participants through the PCT databases enabled research invitations to be sent to a stratified sample, ensuring patients across a range of ages and genders were targeted. Despite the advantages of reaching a specified sample however, this method of recruitment yielded a very poor response rate (n= 23, 4.42%). Conducting the entire research through paper based recruitment would therefore have not been feasible. One possible way of improving the response rate for the mail out recruitment phase may have been to mail a research invitation pack to potential participants asking them to contact the researcher directly in order to obtain an electronic link to the questionnaire sent to their email inbox. This may have enabled a better response rate calculation and been a more user friendly option for potential participants.

Although less useful than response rates in determining if a sample is representative of the true population, survey completion rates were helpful to ascertain the ease of completion and acceptability of the questionnaire. After consenting to take part in the study, 71.1% (n=143) participants completed the entire questionnaire. Those that were mailed the questionnaire had a higher rate of completion (n=47, 77%) compared to those recruited online (n=96, 68.6%).

11.3 Conclusions

Many of the characteristics associated with the web make it a promising resource for public health. For example, the quantity of information available online and ease of access can promote benefits such as public education and empowerment through informed decision-making. The very same characteristics however may lead to negative consequences, such as misinformation or misuse of information (Powell et al., 2011). Research to fully understand the impact of using health-related websites is in its infancy. Research to date has shown some promising results relating to the benefits of using online information resources when compared to traditional forms of information (Wantland et al., 2004). It is important to maximise these benefits, yet information that is considered to be important by patients is often ignored by website developers, for example, experiential information is often poorly addressed (Warren et al., 2014).

In order to advance this area of research, we need to assess the merits and consequences of using various forms of information on health-related websites. Future research needs to assess the delivery of online information in its own right so that the health sector and the health ‘consumer’ are informed in how they can make best use of this resource. Developing a tool which is suitable for use in controlled trials is an essential development if health information is to be systematically and meaningfully evaluated. While some traditional outcome measures may be incorporated into a trial, a focused internet specific instrument may pick up important, yet subtle aspects associated with using information from the web. Using a tool which facilitates the comparison of one website to another (i.e. a control website) also has the advantage of being able to blind participants to the research question, which has been a cause for concern in past research (Samoocha et al., 2010).

This thesis contributes to the growing literature concerning the effects of online information. The primary aim of this thesis was to develop an instrument (the eHealth Impact Questionnaire) to capture the impact of using health-related websites. The questionnaire consists of two independently administered and scored parts (eHIQ-Part 1 and eHIQ-Part 2). The eHIQ-Part 1 consists of 11 items asking about a person's general attitudes towards health-related websites. The eHIQ-Part 2 consists of 26 items asking about a person's views regarding a specific health-related website. Both parts of the questionnaire have a five point response category for all items ranging from 'Strongly disagree to Strongly agree'.

Further research incorporating this instrument may inform future directions in the provision of online information. This standardised tool will enable comparisons between websites which use different approaches in information dissemination and may help inform website developers and healthcare professionals on the benefits of providing additional styles of information (for example, experiential information) to conventional information forms.

11.4 References

ADEMILUYI, G., et al. 2003. Evaluating the reliability and validity of three tools to assess the quality of health information on the Internet. *Patient Education and Counseling,* 50**,** 151-155.

ANDERSON, A. & KLEMM, P. 2008. The Internet: Friend or Foe When Providing Patient Education? *Clinical Journal of Oncology Nursing,* 12**,** 55-63.

ANDRICH, D. 1978. A rating formulation for ordered response categories. *Psychometrika,* 43**,** 561-573.

ANDRICH, D. 1982. An index of person separation in latent trait theory, the traditional KR-20 index, and the Guttman scale response pattern. *Education Research and Perspectives,* 9**,** 95-104.

ANDRICH, D. 1988. *Rasch Models for Measurement. Quantitative applications in the social sciences.,* Thousand Oaks, CA, Sage Publications.

ANDRICH, D. 2004a. Controversy and the Rasch model: a characteristic of incompatible paradigms? *Medical Care,* 42**,** I7-16.

ANDRICH, D. 2004b. Detection of Differential Item Functioning Using Analysis of Variance. *Second International Conference of Measurement in Health Education, Psychology and Marketing: Developments with Rasch Models* Perth, Australia

ANDRICH, D. 2005. Georg Rasch: Mathematician and Statistician. *In:* KEMPF-LEONARD, K. (ed.) *Encyclopedia of Social Measurement.* Amsterdam: Academic Press.

ANDRICH, D. 2011. Rating scales and Rasch measurement. *Expert Review of Pharmacoeconomics and Outcomes Research,* 11**,** 571-85.

ANDRICH, D., et al. 1997. Chapter 4: Diagnostic Opportunities with the Rasch Model for Ordered Response Categories. *In:* ROST, J. & LANGEHEINE, R. (eds.) *Applications of latent trait and latent class models in social sciences.* Michigan: University of Michigan Library.

ATKINSON, L. N., et al. 2009. Using the Internet for Health-Related Activities: Findings From a National Probability Sample. *Journal of Medical Internet Research,* 11**,** e4.

AYANTUNDE, A. A., et al. 2007. A survey of patient satisfaction and use of the internet for health information. *International Journal of Clinical Practice,* 61**,** 458-462.

BABOR, T., et al. 2001. The Alcohol Use Disorders Identification Test, Guidelines for Use in Primary Care. Second ed.

BAKER, R. P., et al. 2004. Development and testing of Web Questionnaires. *In:* PRESSER, S., ROTHGEB, J. M., COUPER, M. P., LESSLER, J. T., MARTIN, E., MARTIN, J. & SINGER, E. (eds.) *Methods for Testing and Evaluating Survey Questionnaires.* Wiley.

BARTLETT, M. S. 1954. A note on multiplying factors for various chi-squared approximations. *Joural of the Royal Statistical Society,* Series B 16**,** 296–298.

BATES, B. R., et al. 2006. The effect of source credibility on consumers' perceptions of the quality of health information on the Internet. *Medical informatics and the Internet in medicine,* 31**,** 45-52.

BAUMER, D., et al. 2014. Advances in motor neurone disease. *J R Soc Med,* 107**,** 14-21.

BEATTY, P. C. & WILLIS, G. B. 2007. Research Synthesis: The Practice of Cognitive Interviewing. *Public Opinion Quarterly,* 71**,** 287-311.

BELVEDERE, S. L. & DE MORTON, N. A. 2010. Application of Rasch analysis in health care is increasing and is applied for variable reasons in mobility instruments. *Journal of Clinical Epidemiology,* 63**,** 1287-1297.

BERNSTAM, E. V., et al. 2005. Instruments to assess the quality of health information on the World Wide Web: what can our patients actually use? *International Journal of Medical Informatics,* 74**,** 13-19.

BETSCH, C., et al. 2011. The influence of narrative v. statistical information on perceiving vaccination risks. *Medical Decision Making,* 31**,** 742-53.

BIDDLE, L., et al. 2008. Suicide and the internet. *British Medical Journal,* 336**,** 800-802.

BLAND, J. & ALTMAN, D. 1995. Multiple significance tests: the Bonferroni method. *British Medical Journal,* 310.

BLANK, G. 2010. *OxIS 2009 Questionnaire* [Online]. Oxford: Oxford Internet Surveys. Available: http://oxis.oii.ox.ac.uk/research/methodology.

BLANK, G. & DUTTON, W. H. 2012. Age and Trust in the Internet: The Centrality of Experience and Attitudes Toward Technology in Britain. *Social Science Computer Review,* 30**,** 135-151.

BOHLMEIJER, E., et al. 2011. Psychometric Properties of the Five Facet Mindfulness Questionnaire in Depressed Adults and Development of a Short Form. *Assessment,* 18**,** 308-320.

BOND, T. G. & FOX, C. M. 2013. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences,* New Jersey, Taylor and Francis.

BOS Bristol Online Surveys, www.survey.bris.ac.uk University of Bristol.

BOWES, P., et al. 2012. 'I need her to be a doctor': patients' experiences of presenting health information from the internet in GP consultations. *British Journal of General Practice* 62.

BOWLING, A. 2005. *Measuring Health: A review of quality of life measurement scales,* Maidenhead, Open University Press.

BOWLING, A. 2009. *Research Methods in Health: Investigating Health and Health Services,* Maidenhead, Open University Press.

BRIGGS, P., et al. 2002. Trust in Online Advice. *Social Science Computer Review,* 20**,** 321-332.

BROD, M., et al. 2009. Qualitative research and content validity: developing best practices based on science and experience. *Quality of Life Research,* 18**,** 1263-1278.

CAIATA-ZUFFEREY, M., et al. 2010. Online Health Information Seeking in the Context of the Medical Consultation in Switzerland. *Qualitative Health Research,* 20**,** 1050-1061.

CATTELL, R. B. 1943. The description of personality: basic traits resolved into clusters. *The Journal of Abnormal and Social Psychology,* 38**,** 476-506.

CATTELL, R. B. 1966. The Scree Test For The Number Of Factors. *Multivariate Behavioral Research,* 1**,** 245-276.

CHAPPLE, A., et al. 2012. An Alarming Prognosis: How People Affected by Pancreatic Cancer Use (and Avoid) Internet Information. *Policy & Internet,* 4**,** 1-20.

CHEW, L., et al. 2004. Brief questions to identify patients with inadequate health literacy. *Family Medicine,* 36**,** 588-594.

CHUNG, J. E. 2013. Social interaction in online support groups: Preference for online social interaction over offline social interaction. *Computers in Human Behavior,* 29**,** 1408-1414.

COATES, J. 2013. *Women, Men and Everyday Talk*, Palgrave Macmillan.

COMPEAU, D. R. & HIGGINS, C. A. 1995. Computer Self-Efficacy: Development of a Measure and Initial Test. *MIS Quarterly,* 19**,** 189-211.

CONWAY, K., et al. 2010. Enhancing Cross-Cultural Appropriateness for Newly Developed Patient-Reported Outcome (PRO) Instruments: the Use of Translatability Assessment. *Patient Reported Outcomes Newsletter* [Online], 44.

CORTI, L. & THOMPSON, P. 2004. Secondary Analysis Of Archived Data *In:* SEALE, C., GOBO, G., GUBRIUM, J. F. & SILVERMAN, D. (eds.) *Qualitative Research Practice* London, England: SAGE Publications Ltd.

COURTENAY-QUIRK, C., et al. 2010. Perceptions of HIV-related websites among persons recently diagnosed with HIV. *AIDS Patient Care STDS,* 24**,** 105-15.

CURTIS, D. D. 2004. Person misfit in attitude surveys: influences, impacts and implications. *International Education Journal,* 5**,** 125-143.

DARGENTAS, M. & ROUX, D. L. 2005. Potentials and Limits of Secondary Analysis in a Specific Applied Context: The Case of EDF—Verbatim. *Forum: Qualitiative Social Research,* 6.

DAVIS, F. D. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly,* 13**,** 319-340.

DE BRUIN, A. F., et al. 1994. The development of a short generic version of the sickness impact profile. *Journal of Clinical Epidemiology,* 47**,** 407-418.

DEVAUS, D. A. 2002. *Surveys in social research,* Crows Nest, Australia, Allen and Unwin.

DEVELLIS, R. F. 2006. Classical Test Theory. *Medical Care,* 44**,** 50-59

DEWALT, D. A., et al. 2007. Evaluation of Item Candidates: The PROMIS Qualitative Item Review. *Medical Care,* 45**,** 12-21.

DILLMAN, D. A. 2011. *Mail and Internet Surveys: The Tailored Design Method,* New Jersey, John Wiley and Sons.

DILLMAN, D. A. & BOWKER, D. K. 2001. The Web Questionnaire Challenge to Survey Methodologists. *In:* REIPS, U. D. & BOSNJAK, M. (eds.) *Dimensions of Internet Science.* Lengerich: Pabst Science Publishers.

DIPEX. 2014. *Database of Individual Patient Experiences* [Online]. Available: www.healthtalkonline.org.

DOBROW, J. M., et al. 2008. Response Audit of an Internet Survey of Health Care Providers and Administrators: Implications for Determination of Response Rates. *Journal of Medical Internet Research,* 10**,** e30.

DUBENSKE, L. L., et al. 2009. Psychometric Evaluation of the Health Information Orientation Scale. *Journal of Health Psychology,* 14**,** 721-730.

DUTTA-BERGMAN, M. 2003. Trusted Online Sources of Health Information: Differences in Demographics, Health Beliefs, and Health-Information Orientation. *Journal of Medical Internet Research,* 5**,** e21.

DUTTA-BERGMAN, M. 2004. Health Attitudes, Health Cognitions, and Health Behaviors among Internet Health Information Seekers: Population-Based Survey. *Journal of Medical Internet Research,* 6.

DUTTON, W. H. & BLANK, G. 2011. Next Generation Users: The Internet in Britain. *Oxford Internet Survey 2011 Report.* Oxford Internet Institute, University of Oxford.

DUTTON, W. H., et al. 2013. Cultures of the Internet:The Internet in Britain. *Oxford Internet Survey 2013 Report.* Oxford Internet Institute, University of Oxford.

DUTTON, W. H. & SHEPHERD, A. 2006. Trust in the Internet as an experience technology. *Information, Communication & Society,* 9**,** 433-451.

EMA 2004. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. European Medicine Agency.

EMBRETSON, S. E. & REISE, S. P. 2000. *Item Response Theory*, Taylor & Francis.

ENTWISTLE, V. A., et al. 2011. How information about other people's personal experiences can help with healthcare decision-making: A qualitative study. *Patient Education and Counseling,* 85**,** e291-e298.

ESRC. 2012. *New national digital repository for social and economic data* [Online]. Economic and Social Research Council. Available: http://www.esrc.ac.uk/news-and-events/press-releases/22231/new-national-digital-repository-for-social-and-economic-data.aspx [Accessed 22/01/2013].

ETTER, J.-F. & PERNEGER, T. V. 2001. A comparison of cigarette smokers recruited through the Internet or by mail. *International Journal of Epidemiology,* 30**,** 521-525.

EUROSTAT 23.01.2014. Individuals using the Internet for seeking health-related information *Information Society Statistics, Computers and the Internet in households and enterprises, Internet - Level of access, use and activities* European Commission.

EUROSTAT 31.07.2013. Households having access to the Internet, by type of connection - % of all households. *Information Society Statistics, Computers and the Internet in households and enterprises, Households having access to the Internet, by type of connection.* European Commission.

EVANS, V. P. 1999. Higher-Order Factor Analysis: An Introductory Primer. *Annual Meeting of the Southwest Educational Research Association 1999.*

EYSENBACH, G. 2004. Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *Journal of Medical Internet Research,* 6**,** e34.

EYSENBACH, G. 2008. Medicine 2.0: Social Networking, Collaboration, Participation, Apomediation, and Openness. *Journal of Medical Internet Research,* 10**,** e22.

EYSENBACH, G., et al. 2002. Empirical Studies Assessing the Quality of Health Information for Consumers on the World Wide Web. *JAMA: The Journal of the American Medical Association,* 287**,** 2691-2700.

EYSENBACH, G. & WYATT, J. 2002. Using the Internet for Surveys and Health Research. *Journal of Medical Internet Research,* 4**,** e13.

FDA 2009. US Department of Health and Human Services Food and Drug Administration. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims.

FISHER, W. 1992. Reliability statistics. *Rasch Measurement Transactions,* 6**,** 238.

FOX, S. 2011a. Health Topics. Washington, DC: Pew Internet and American Life Project.

FOX, S. 2011b. Peer-to-peer healthcare. Washington, DC: Pew Internet and American Life Project.

FOX, S. & DUGGAN, M. 2013. Health Online 2013. Washington, DC: Pew Internet and American Life Project.

FOX, S., et al. 2000. The online health care revolution: How the Web helps Americans take better care of themselves. *The Pew Internet & American Life Project.* Washington.

FRIEDMAN, H. H. & AMOO, T. 1999. Rating the rating scales. *Journal of Marketing Management,* 9**,** 114-123.

FROST, H. J. & MASSAGLI, P. M. 2008. Social Uses of Personal Health Information Within PatientsLikeMe, an Online Patient Community: What Can Happen When Patients Have Access to One Another's Data. *Journal of Medical Internet Research,* 10**,** e15.

GORSUCH, R. L. 1983. *Factor Analysis*, L. Erlbaum Associates.

GRAHAM, L. A., et al. 2008. Online Advertising as a Public Health and Recruitment Tool: Comparison of Different Media Campaigns to Increase Demand for Smoking Cessation Interventions. *Journal of Medical Internet Research,* 10**,** e50.

GUTTMAN, L. 1944. A Basis for Scaling Qualitative Data. *American Sociological Review,* 9**,** 139-150.

HAMBLETON, R. K. & JONES, R. W. 1993. An NCME Instructional Module on. *Educational Measurement: Issues and Practice,* 12**,** 38-47.

HANIF, F., et al. 2009. The role of quality tools in assessing reliability of the Internet for health information. *Informatics for Health and Social Care,* 34**,** 231-243.

HARRIS, P. R., et al. 2009. The effect of credibility-related design cues on responses to a web-based message about the breast cancer risks from alcohol: randomized controlled trial. *Journal of Medical Internet Research,* 11.

HARRIS, P. R., et al. 2011. Perceived Threat and Corroboration: Key Factors That Improve a Predictive Model of Trust in Internet-based Health Information and Advice. *Journal of Medical Internet Research,* 13**,** e51.

HAYS, R. D. & REVICKI, D. 2005. Reliability and validity (including responsiveness). *In:* FAYERS, P. M. & HAYS, R. D. (eds.) *Assessing Quality Of Life In Clinical Trials: Methods And Practice.* Oxford University Press.

HEATON, J. 1998. Secondary analysis of qualitative data. *Sociology at Surrey,* Social Research Update 22.

HEATON, J. 2004. *Reworking Qualitative Date* London, Sage Publications.

HEINEMANN, A. W., et al. 1993. Relationships between impairment and physical disability as measured by the functional independence measure. *Arch Phys Med Rehabil,* 74**,** 566-73.

HELMSTADTER, G. C. 1964. *Principles of psychological measurement,* New York, Appleton-Century-Crofts.

HIBBARD, J. H., et al. 2005. Development and Testing of a Short Form of the Patient Activation Measure. *Health Services Research,* 40**,** 1918-1930.

HIBBARD, J. H., et al. 2004. Development of the Patient Activation Measure (PAM): Conceptualizing and Measuring Activation in Patients and Consumers. *Health Services Research,* 39**,** 1005-1026.

HIGGINS, et al. 2011. A literature review on health information seeking behaviour on the web: a health consumer and health professional perspective. Stockholm: ECDC.

HINKIN, T. R. 1998. A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires. *Organizational Research Methods,* 1**,** 104-121.

HINTON, L., et al. 2010. Infertility; isolation and the Internet: A qualitative interview study. *Patient Education and Counseling,* 81**,** 436-441.

HOWIE, J. G., et al. 1998. A comparison of a Patient Enablement Instrument (PEI) against two established satisfaction scales as an outcome measure of primary care consultations. *Family Practice,* 15**,** 165-171.

HU, X., et al. 2012. The Prepared Patient: Information Seeking of Online Support Group Members Before Their Medical Appointments. *Journal of Health Communication,* 17**,** 960-978.

HUNT, K., et al. 2009. `My brain couldn't move from planning a birth to planning a funeral': A qualitative study of parents' experiences of decisions after ending a pregnancy for fetal abnormality. *International Journal of Nursing Studies,* 46**,** 1111-1121.

IBM 2011. SPSS Statistics for Windows. *Version 20.0.* Armonk, NY: IBM Corp. .

JAMIESON, S. 2004. Likert scales: how to (ab)use them. *Medical Education,* 38**,** 1217-1218.

JENKINSON, C. 1997. Assessment and Evaluation of Health Care and Medical Care: an introduction and overview. *In:* JENKINSON, C. (ed.) *Assessment and Evaluation of Health Care and Medical Care, A methods Text.* Maidenhead, Philadelphia: Open University Press.

JENKINSON, C., et al. 2013. The Stroke Impact Scale: Validation in a UK Setting and Development of a SIS Short Form and SIS Index. *Stroke,* 44**,** 2532-2535.

JENKINSON, C. & MCGEE, H. 1997. Patient assessed outcomes: measuring health status and quality of life. *In:* JENKINSON, C. (ed.) *Assessment and evaluation of health and medical care: a methods text.* Oxford: Open University Press.

JENKINSON, C., et al. 1994. Measuring change over time: a comparison of results from a global single item of health status and the multi-dimensional SF-36 health status survey questionnaire in patients presenting with menorrhagia. *Quality of Life Research,* 3**,** 317-321.

JONES, R. 2013. Development of a Questionnaire and Cross-Sectional Survey of Patient eHealth Readiness and eHealth Inequalities. *Medicine 2.0,* 2**,** e9.

KAISER, H. F. & RICE, J. 1974. Little Jiffy, Mark Iv. *Educational and Psychological Measurement,* 34**,** 111-117.

KERR, C., et al. 2010. Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcomes research. *Expert Review of Pharmacoeconomics & Outcomes Research,* 10**,** 269-281.

KIM, N. E., et al. 2012. The impact of user's perceived ability on online health information acceptance. *Telemedicine Journal and E-Health,* 18**,** 703-8.

KLINE, P. 1994. *An easy guide to factor analysis*, Routledge, Chapman & Hall, Incorporated.

KLINE, P. 2000. *A Psychometric Primer,* London, Free Association Books.

KOO, M. & SKINNER, H. 2005. Challenges of Internet Recruitment: A Case Study with Disappointing Results. *Journal of Medical Internet Research,* 7**,** e6.

KUMMERVOLD, E. P., et al. 2008. eHealth Trends in Europe 2005-2007: A Population-Based Survey. *Journal of Medical Internet Research,* 10**,** e42.

LAGAN, B. M., et al. 2010. Internet Use in Pregnancy Informs Women's Decision Making: A Web-Based Survey. *Birth,* 37**,** 106-115.

LINACRE, J. M. 2010. When to stop removing items and persons in Rasch misfit analysis? *Rasch Measurement Transactions,* 23**,** 1241.

LOHR, K. N. 2002. Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research,* 11**,** 193-205.

LOWE, P., et al. 2009. "Making it All Normal": The Role of the Internet in Problematic Pregnancy. *Qualitative Health Research,* 19**,** 1476-1484.

LUSTRIA, M. L. A., et al. 2009. Computer-tailored health interventions delivered over the web: Review and analysis of key components. *Patient Education and Counseling,* 74**,** 156-173.

MARRIE, A. R., et al. 2013. Preferred Sources of Health Information in Persons With Multiple Sclerosis: Degree of Trust and Information Sought. *Journal of Medical Internet Research,* 15**,** e67.

MASTERS, G. 1982. A rasch model for partial credit scoring. *Psychometrika,* 47**,** 149-174.

MATSUNAGA, M. 2010. How to Factor-Analyze Your Data Right: Do's, Don'ts, and How-To's. *International Journal of Psychological research,* 3**,** 97-110.

MAZANDERANI, F., et al. 2012. Being differently the same: the mediation of identity tensions in the sharing of illness experiences. *Social Science and Medicine,* 74**,** 546-53.

MCCOLL, E. 2005. Developing questionnaires. *In:* FAYERS, P. & HAYS, R. (eds.) *Assessing Quality of Life in Clinical Trials* Second ed. Oxford: Oxford Medical Publications

MCHORNEY, C. A. & MONAHAN, P. O. 2004. Postscript: Applications of Rasch Analysis in Health Care. *Medical Care,* 42**,** I73-I78.

MCMULLAN, M. 2006. Patients using the Internet to obtain health information: How this affects the patient–health professional relationship. *Patient Education and Counseling,* 63**,** 24-28.

MESSICK, S. 1989. Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educational Researcher,* 18**,** 5-11.

MO, P. K. H. & COULSON, N. S. 2010. Empowering processes in online support groups among people living with HIV/AIDS: A comparative analysis of 'lurkers' and 'posters'. *Computers in Human Behaviour,* 26**,** 1183-1193.

MORGAN, J. A., et al. 2013. Internet-Based Recruitment to a Depression Prevention Intervention: Lessons From the Mood Memos Study. *Journal of Medical Internet Research,* 15**,** e31.

MORLEY, D., et al. 2013. The PDQ-Carer: Development and validation of a summary index score. *Parkinsonism & Related Disorders,* 19**,** 448-449.

MURRAY, E., et al. 2005. Interactive Health Communication Applications for people with chronic disease. *Cochrane Database Systematic Review,* 19.

MURRAY, E., et al. 2009. Methodological Challenges in Online Trials. *Journal of Medical Internet Research,* 11**,** e9.

NATCEN. *Framework* [Online]. Available: http://www.natcen.ac.uk/about-us/our-approach/framework [Accessed 30.09.12.

NCB 2010. Chapter 5: Online health information. *Medical profiling and online medicine: the ethics of 'personalised healthcare' in a consumer age.* Nuffield Council on Bioethics.

NELSON, D., et al. 2004. The Health Information National Trends Survey (HINTS): Development, Design, and Dissemination. *Journal of Health Communication,* 9**,** 443-460.

NHS_ENGLAND. 2013a. *The GP Patient Survey: July 2012 – March 2013 Summary Report* [Online]. Available: http://www.gp-patient.co.uk/results/ [Accessed 20.01.14.

NHS_ENGLAND. 2013b. *The Information Standard* [Online]. Available: http://www.theinformationstandard.org/ [Accessed 21.02.14.

NORMAN, D. C. & SKINNER, A. H. 2006. eHEALS: The eHealth Literacy Scale. *Journal of Medical Internet Research,* 8**,** e27.

NORMAN, G. & STREINER, D. 2000. *Biostatistics: The Bare Essentials* London, B.C. Decker.

NUNNALLY, J. & BERNSTEIN, I. H. 1994. *Psychometric Theory*, New York: McGraw-Hill.

O'NEILL, B., et al. 2012. People power or pester power? Multiple sclerosis, YouTube, and the liberation procedure. *The Lancet,* 380**,** S20.

O'NEILL, B., et al. 2014. User-generated online health content: a survey of Internet users in the United kingdom. *Journal of Medical Internet Research,* 16.

OFCOM 2013. The Communications Market 2013. *Internet and web-based content, UK CMR internet.*

OLSON, K. 2010. An Examination of Questionnaire Evaluation by Expert Reviewers. *Field Methods,* 22**,** 295-318.

ONS 2013. Office for National Statistics, Internet Access - Households and Individuals, 2013. Statistical Bulletin.

ONS 2013 Table 1A. Office for National Statistics, Internet Access Quarterly Update, Statistical Bulletin.

ONS 2013, Table 1B. Office for National Statistics, Internet Access Quarterly Update, Statistical Bulletin.

ONS 2013, Table 5B. Office for National Statistics, Internet Access Quarterly Update, Statistical Bulletin.

PALLANT, J. F. & TENNANT, A. 2007. An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology,* 46**,** 1-18.

PAWSON, R. 2006. *Evidence-Based Policy: A Realist Perspective,* London, Sage Publications.

PAWSON, R., et al. 2005. Realist review – a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy,* 10**,** 21-34.

PETT, M. A., et al. 2003. *Making Sense of Factor Analysis: The Use of Factor Analysis for Instrument Development in Health Care Research*, SAGE Publications.

POPE, C., et al. 2000. Analysing qualitative data. *British Medical Journal,* 320**,** 114-116.

PORTNEY, L. & WATKINS, M. 2009. *Foundations of Clinical Research Applications to practice*, Pearson Education: New Jersey.

POTOSKY, D. 2007. The Internet knowledge (iKnow) measure. *Computers in Human Behavior,* 23**,** 2760-2777.

POWELL, J., et al. 2012. Experience of a Health Website Evaluated in a Randomised Trial (EXPERT). *Trial registration number: ISRCTN29549695* DOI 10.1186/ISRCTN29549695

POWELL, J., et al. 2011. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. *Journal of Medical Internet Research,* 13.

POWELL, J. A., et al. 2005. A critical analysis of the literature on the Internet and consumer health information. *Journal of Telemedicine and Telecare,* 1**,** 41-3.

PRAINSACK, B. 2013. Let's get real about virtual: online health is here to stay. *Genetics Research,* 95**,** 111-113.

PULMAN, A. 2010. A patient centred framework for improving LTC quality of life through Web 2.0 technology. *Health Informatics Journal,* 16**,** 15-23.

QSR_INTERNATIONAL 2010. NVivo qualitative data analysis software. Version 9 ed.: QSR International Pty Ltd.

QUALTRICS 2005. Qualtrics. Version 37,892 ed. Provo, Utah, USA.: Qualtrics Research Suite.

RAINIE, L. & SMITH, A. 2013. Tablet and E-reader Ownership Update. DC, Wasington.

RAINS, S. A. 2007. Perceptions of traditional information sources and use of the world wide web to seek health information: findings from the health information national trends survey. *Journal of Health Communication,* 12**,** 667-80.

RASCH, G. 1960. *Probabilistic models for some intelligence and attainment tests. ,* Chicago, University of Chicago.

REEVE, B., et al. 2013. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Quality of Life Research***,** 1-17.

REIPS, U. D. 2002. Standards for Internet-based experimenting. *Experimental Psychology,* 49**,** 243-256.

REIS, S., et al. 2013. Health information and communication technology in healthcare communication: The good, the bad, and the transformative. *Patient Education and Counseling,* 93**,** 359-362.

RING, L., et al. 2010. Putting the text back into context: toward increased use of mixed methods for quality of life research. *Quality of Life Research,* 19**,** 613-615.

RITCHIE, J. & SPENCER, L. 1994. Qualitative data analysis for applied policy research. *In:* BRYMAN, A. & BURGESS, R. G. (eds.) *Analysing Qualitative Data* London: Routledge.

ROBSON, C. 2002. *Real world research: a resource for social scientists and practitioner-researchers,* Chichester, UK, Blackwell Publishers.

ROCK, M. 2010. Harnessing Social Networks along with Consumer-Driven Electronic Communication Technologies to Identify and Engage Members of 'Hard-to-Reach' Populations: A Methodological Case Report. *BMC Medical Research Methodology,* 10**,** 1-4.

ROULEAU, C. R. & VON RANSON, K. M. 2011. Potential risks of pro-eating disorder websites. *Clinical Psychology Review,* 31**,** 525-531.

RUBIN, D. B. 1976. Inference and Missing Data. *Biometrika,* 63**,** 581-592.

RYAN, J. P. 1983. Introduction to latent trait analysis and item response theory. *In:* HATHAWAY, W. E. (ed.) *Testing in the Schools. New Directions for Testing and Measurement.* San Francisco: Jossey-Bass.

SAMOOCHA, D., et al. 2010. Effectiveness of web-based interventions on patient empowerment: a systematic review and meta-analysis. *Journal of Medical Internet Research,* 12.

SEÇKIN, G. 2011. Informational and decisional empowerment in online health support communities: initial psychometric validation of the Cyber Info-Decisional Empowerment Scale (CIDES) and preliminary data from administration of the scale. *Supportive Care in Cancer,* 19**,** 2057-2061.

SETHURAM, R. & WEERAKKODY, A. N. A. 2010. Health information on the internet. *Journal of Obstetrics & Gynaecology,* 30**,** 119-121.

SIEGEL, E. R., et al. 2009. Web-based public health information dissemination and evaluation. *In:* DETELS, R., BEAGLEHOLE, R., LANSANG, M. A. & GULLIFORD, M. (eds.) *Oxford Textbook of Public Health.* Fifth ed. New York: Oxford University Press.

SILLENCE, E. & BRIGGS, P. 2009. Examining the role of the Internet in health behaviour. *In:* JOINSON, A., MCKENNA, K. & POSTMES, T. (eds.) *Oxford Handbook of Internet Psychology.* Oxford: Oxford Uuniveristy Press.

SILLENCE, E., et al. 2004. Trust and mistrust of online health sites. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* Vienna, Austria: ACM.

SILLENCE, E., et al. 2006. A framework for understanding trust factors in web-based health advice. *International Journal of Human-Computer Studies,* 64**,** 697-713.

SILLENCE, E., et al. 2007a. Going online for health advice: Changes in usage and trust practices over the last five years. *Interacting with Computers,* 19**,** 397-406.

SILLENCE, E., et al. 2007b. How do patients evaluate and make use of online health information? *Social Science & Medicine,* 64**,** 1853-1862.

SILVERSTEIN, B., et al. 1991. Applying psychometric criteria to functional assessment in medical rehabilitation: I. Exploring unidimensionality. *Arch Phys Med Rehabil,* 72**,** 631-7.

SLEGERS, K., et al. 2008. Effects of Computer Training and Internet Usage on the Well-Being and Quality of Life of Older Adults: A Randomized, Controlled Study. *The*

Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 63**,** 176-P184.

SMITH, E. V., JR. 2002. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement,* 3**,** 205-31.

SPIELBERGER CD, et al. 1989. *State-Trait Anxiety Inventory: A Comprehensive Bibliography,* Palo Alto, CA, Consulting Psychologists Press.

STELLEFSON, M., et al. 2013. Web 2.0 Chronic Disease Self-Management for Older Adults: A Systematic Review. *Journal of Medical Internet Research,* 15**,** e35.

STREINER, D. & NORMAN, G. 2008. *Health Measurement Scales: A practical guide to their development and use,* Oxford, Oxford University Press.

SUE, V. M. & RITTER, L. A. 2012. *Conducting Online Surveys,* Thousand Oaks, California, Sage Publications.

TENNANT, A. & CONAGHAN, P. G. 2007. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care & Research,* 57**,** 1358-1362.

TENNANT, A., et al. 2004. Application of Rasch analysis in the development and application of quality of life instruments. *Value Health,* 7**,** S22-6.

TESIO, L. 2003. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. *J Rehabil Med,* 35**,** 105-15.

THOMSON_REUTERS 1998-2010. Endnote Version X4.0.2 (Bld 5149) ed.

THURSTONE, L. L. 1928. Attitudes can be measured. *American Journal of Sociology,* 33**,** 529-554.

TOEPOEL, V., et al. 2009. Design of Web Questionnaires. *Sociological Methods & Research,* 37**,** 371-392.

VAN UDEN-KRAAN, C. F., et al. 2008. Empowering Processes and Outcomes of Participation in Online Support Groups for Patients With Breast Cancer, Arthritis, or Fibromyalgia. *Qualitative Health Research,* 18**,** 405-417.

WAGNER, L. S. & WAGNER, T. H. 2003. The Effect of Age on the Use of Health and Self-Care Information: Confronting the Stereotype. *The Gerontologist,* 43**,** 318-324.

WALLACE, L., et al. 2006. Brief report: screening items to identify patients with limited health literacy skills. *The Journal of General Internal Medicine,* 21**,** 874-877.

WALLACE, P., et al. 2011. On-line Randomized Controlled Trial of an Internet Based Psychologically Enhanced Intervention for People with Hazardous Alcohol Consumption. *PLOS ONE,* 6**,** e14740.

WANTLAND, J. D., et al. 2004. The Effectiveness of Web-Based vs. Non-Web-Based Interventions: A Meta-Analysis of Behavioral Change Outcomes. *Journal of Medical Internet Research,* 6**,** e40.

WARE, J., et al. 1994. *SF-36 physical and mental summary scores: a user's manual.,* Boston, The Health Institute, New England Medical center.

WARE, J. E., JR., et al. 1995. Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. *Medical Care,* 33**,** AS264-79.

WARE, J. E., JR. & SHERBOURNE, C. D. 1992. The MOS 36-item short-form health survey (SF-36). Conceptual framework and item selection. *Medical Care,* 30**,** 473-83.

WARREN, E., et al. 2014. Do cancer-specific websites meet patient's information needs? *Patient Education and Counselling,* 95**,** 126-36.

WHO 2008. WHO STEPS Instrument (Core and Expanded). *The WHO STEPwise approach to chronic disease risk factor surveillance (STEPS).* Geneva: World Health Organization.

WICKS, P., et al. 2010. Sharing health data for better outcomes on PatientsLikeMe. *Journal of Medical Internet Research,* 12.

WILLIAMS, D. 2006. On and Off the 'Net: Scales for Social Capital in an Online Era *Journal of Computer-Mediated Communication,* 11.

WILLIAMS, G. 2009. Statistical methods. *In:* DETELS, R., BEAGLEHOLE, R., LANSANG, M. A. & GULLIFORD, M. (eds.) *Oxford Textbook of Public Health* Fifth ed. New York: Oxford University Press.

WILLIS, G. B. 1999. Cognitive Interviewing: A “How To” Guide. *Short course presented at the 1999 Meeting of the American Statistical Association.* Research Triangle Institute.

WILLIS, G. B. 2005a. *Cognitive Interviewing in Practice: Think-Aloud, Verbal Probing, and Other Techniques. Cognitive Interviewing. ,* Thousand Oaks, USA, Sage Publications.

WILLIS, G. B. 2005b. *Cognitive interviewing: a tool for improving questionnaire design,* London, Sage Publications.

WILLIS, G. B. & LESSLER, J. T. August, 1999. Question Appraisal System (QAS-99). Rockville: Research Triangle Institute.

WILSON, M. 2005. *Constructing measures: an item response modeling approach,* New Jersey, Lawrence Erlbaum Associates.

WILSON, P. & RISK, A. 2002. How to find the good and avoid the bad or ugly: a short guide to tools for rating quality of health information on the internetCommentary: On the way to quality. *British Medical Journal,* 324**,** 598-602.

WRIGHT, B. 1980. Foreward and Introduction. *Probabilistic models for some intelligence and attainment tests. G Rasch.* Chicago The University of Chicago Press.

WRIGHT, B. D. & MASTERS, G. N. 1982. *Rating Scale Analysis. Rasch Measurement,* Chicago, Pluribus Press.

WRIGHT, B. D. & STONE, M. H. 1979. The Measurement Model. *Best test design.* Chicago: Mesa Press.

WYKE, J. V. & ANDRICH, D. 2006. A typology of polytomously scored mathematics items disclosed by the Rasch model: implications for constructing a continuum of achievement. *ARC Linkage Report 2.* Murdoch University, Western Australia.

WYKE, S., et al. 2011. Information for choice: what people need, prefer and use - Report for the National Institute for Health Research Service Delivery and Organisation programme. *National Institute for Health Research. SDO Project, 08/1710/153. NIHR Service Delivery and Organisation (SDO).* Southampton.

YARDLEY, L., et al. 2010. Evaluation of a Web-based intervention providing tailored advice for self-management of minor respiratory symptoms: exploratory randomized controlled trial. *Journal of Medical Internet Research,* 12.

YARDLEY, L., et al. 2011. Using Mixed Methods to Design a Web-based Behavioural Intervention to Reduce Transmission of Colds and Flu. *Journal of Health Psychology,* 16**,** 353-364.

YBARRA, M. L. & SUMAN, M. 2006. Help seeking behavior and the Internet: a national survey. *International Journal of Medical Informatics,* 75**,** 29-41.

ZIEBLAND, S. & HUNT, K. 2014. Using secondary analysis of qualitative data of patient experiences of health care to inform health services research and policy. *Journal of Health Services Research & Policy,* 26**,** 26.

ZIEBLAND, S. & MCPHERSON, A. 2006. Making sense of qualitative data analysis: an introduction with illustrations from DIPEx (personal experiences of health and illness). *Medical Education,* 40**,** 405-414.

ZIEBLAND, S. & WRIGHT, L. 1997. Qualitative Research Methods. *In:* JENKINSON, C. (ed.) *Assessment and Evaluation of Health Care and Medical Care.* Maidenhead: Open University Press.

ZIEBLAND, S. & WYKE, S. 2012. Health and illness in a connected world: how might seeing and sharing experiences on the internet affect people’s health? . *The Milbank Quarterly,* 90**,** 219-49.