**Applying digital technology to the prediction of depression and anxiety in older adults**

Jacob Andrews

Centre for Assistive Technology and Connected Healthcare

School of Health and Related Research

The University of Sheffield

**A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy**

**Acknowledgements**

I would like to express my gratitude firstly to all members of my supervisory team, who have patiently guided me through all stages of the PhD, and spent hours reviewing my work and adding comments to help me to improve it. Their input has taught me a huge amount and I am very grateful.

I would also like to thank all the research participants who generously gave up their time to take part in the research. I would like to thank Kathryn, Bridget and John at Sheffield 50+, as well as Janet at Roman Ridge, Lauren in CATCH, and divisions of the University of the Third Age across the UK. I’d also like to thank Jenni and Rachel for their help in organising the final study.

For all of their moral support, I’d like to thank my mother, my father, my brother, Emma, all of my fellow PhD students in ScHARR (especially Matt, Dave and Phil) and Dorothy, who all had their part to play at different times. George and Dave have also been rocks I have depended upon in this process. I’d also like to thank my friends at the Sheffield Buddhist Centre, who have supported me with their kind words.

I am also grateful to Sarah S, Kat, Stephen and Sarah H in CATCH for their guidance and support.

I would finally like to thank Research and Innovation Services at the University of Sheffield for funding my PhD.

**Abbreviations used in this thesis**

4GDS 4-item Geriatric Depression Scale

6CIT 6-item Cognitive Impairment Test

ACTS Accelerated Creation to Sustainment

AUC Area Under the Curve

AUROC Area Under the Receiver Operating Characteristic curve

CATCH Centre for Assistive Technology and Connected Healthcare

CBT Cognitive Behavioural Therapy

cCBT Computerised Cognitive Behavioural Therapy

COBALT Challenging Obstacles and Barriers to Assistive Living Technologies

COPD Chronic Obstructive Pulmonary Disease

DSM-5 Diagnostic and Statistical Manual of Mental Disorders 5

GAD7 General Anxiety Disorder 7

GAI Geriatric Anxiety Inventory

GDS Geriatric Depression Scale

HADS Hospital Anxiety and Depression Scale

ICBT Internet Cognitive Behavioural Therapy

ICT Integrated Care Team

NANA Novel Assessment of Nutrition in Ageing

LASSO Least Absolute Shrinkage and Selection Operator

NHS National Health Service

NICE National Institute for Health and Care Excellence

PHQ-9 Patient Health Questionnaire 9

PTSD Post-Traumatic Stress Disorder

ROC Receiver Operating Characteristic

ScHARR School of Health and Related Research

SSRI Selective Serotonin Re-uptake Inhibitor

U3A University of the Third Age

WHO World Health Organisation

**Journal articles and conference presentations arising from the work presented in this thesis**

**Journal articles**

Andrews, J.A., Harrison, R.F., Brown, L.J.E., MacLean, L.M., Hwang, F., Smith, T., Williams, E.A., Timon, C., Adlam, T., Khadra, H. and Astell A.J., (2017). Using the NANA toolkit at home to predict older adults' future depression. Journal of Affective Disorders. Vol. 213, pp. 187-190.

Andrews, J.A., Brown, L.J.E., Astell, A.J. (Submitted) An exploration of touchscreen technology for older adults to research and report on mental health. *Journal of Medical Internet Research*.

**Conference proceedings**

Andrews J.A., Astell A.J., Brown L.J.E., Harrison R.F. & Hawley M.S. (2017) Technology for Early Detection of Depression and Anxiety in Older People. Studies in Health Technology and Informatics, 242, 374-380.

**Conference abstracts**

Andrews, J.A., Harrison, R.F., Brown, L.J.E., Astell, A.J. (2015). Developing a model for detection of the onset of depression in older adults. Canadian Association on Gerontology Conference 2015.

Andrews, J.A., Astell, A.J., Brown, L.J.E., (2016). Technology for mental health in later life: Is it a practical solution? British Society of Gerontology Conference 2016.

Andrews, J.A., Harrison, R.F., Brown, L.J.E., Hawley, M.S. (2017). Predicting anxiety in older adults using self-reported mood and appetite data from digital devices*. International Society of Research on Internet Interventions 9th Scientific Meeting*.

# **Table of Contents**

[Table of Figures 10](#_Toc513973663)

[Table of Tables 12](#_Toc513973664)

[Abstract 13](#_Toc513973665)

[Chapter 1 – Introduction 15](#_Toc513973666)

[1.1 Introduction 15](#_Toc513973667)

[1.2 Rationale 15](#_Toc513973668)

[Depression and anxiety in older people 15](#_Toc513973669)

[Early intervention 16](#_Toc513973670)

[Current practices in the detection of depression and anxiety 18](#_Toc513973671)

[Digital technologies and older adults 19](#_Toc513973672)

[Machine learning 21](#_Toc513973673)

[Implementation of digital technologies within healthcare 22](#_Toc513973674)

[Summary 22](#_Toc513973675)

[1.3 Aim and objectives 23](#_Toc513973676)

[1.4 Summary of the research completed 24](#_Toc513973677)

[Chapter 2 – Narrative literature review 30](#_Toc513973678)

[2.1 Introduction 30](#_Toc513973679)

[2.2 Method 31](#_Toc513973680)

[Approach 31](#_Toc513973681)

[Process 31](#_Toc513973682)

[2.3 Depression and anxiety in older adults 32](#_Toc513973683)

[Depression and its prevalence in older adults 32](#_Toc513973684)

[Anxiety and its prevalence in older adults 34](#_Toc513973685)

[Treatments for depression and anxiety 34](#_Toc513973686)

[Risk factors and correlates of depression and anxiety in later life 35](#_Toc513973687)

[2.4 Detection and diagnosis of depression and anxiety in older adults 36](#_Toc513973688)

[Symptom identification and diagnosis 36](#_Toc513973689)

[Barriers to detection and diagnosis 37](#_Toc513973690)

[2.5 Mental health technology 39](#_Toc513973691)

[2.6 Older adults and technology 41](#_Toc513973692)

[2.7 Implementation of technology in healthcare 45](#_Toc513973693)

[Chapter 3 (A) – Systematic Mapping Review 1 50](#_Toc513973694)

[3.1 Introduction 50](#_Toc513973695)

[Aims 50](#_Toc513973696)

[Research Question 51](#_Toc513973697)

[3.2 Methods 51](#_Toc513973698)

[Decision of search terms to include 51](#_Toc513973699)

[Final search terms 52](#_Toc513973700)

[Databases searched 53](#_Toc513973701)

[Inclusion/exclusion criteria 54](#_Toc513973702)

[Search technique 54](#_Toc513973703)

[Sifting 54](#_Toc513973704)

[3.3 Results 55](#_Toc513973705)

[Title sift 56](#_Toc513973706)

[Abstract sift 57](#_Toc513973707)

[Full paper sift 58](#_Toc513973708)

[Information Extraction 58](#_Toc513973709)

[Characterisation of studies included 66](#_Toc513973710)

[3.4 Discussion 69](#_Toc513973711)

[Limitations 72](#_Toc513973712)

[Conclusion 72](#_Toc513973713)

[Chapter 3 (B) Systematic Mapping Review 2 73](#_Toc513973714)

[3.5 Introduction 73](#_Toc513973715)

[Background 73](#_Toc513973716)

[Aims 73](#_Toc513973717)

[Research question 73](#_Toc513973718)

[3.6 Methods 74](#_Toc513973719)

[Decision of search terms to include 74](#_Toc513973720)

[Final search terms 74](#_Toc513973721)

[Databases searched 75](#_Toc513973722)

[Inclusion criteria 75](#_Toc513973723)

[Search technique 76](#_Toc513973724)

[Sifting 77](#_Toc513973725)

[Data extraction 77](#_Toc513973726)

[3.7 Results 77](#_Toc513973727)

[Characterisation of studies included 81](#_Toc513973728)

[3.8 Discussion 83](#_Toc513973729)

[Limitations 85](#_Toc513973730)

[Conclusion 85](#_Toc513973731)

[Chapter 4 – An exploratory machine learning approach to the prediction of future depression 87](#_Toc513973732)

[4.1 Introduction 87](#_Toc513973733)

[4.1.1 Rationale 87](#_Toc513973734)

[4.1.2 Differentiating machine learning and statistical inference 88](#_Toc513973735)

[4.1.3 Summary 91](#_Toc513973736)

[4.1.4 Aims 91](#_Toc513973737)

[4.2 The NANA dataset 92](#_Toc513973738)

[4.2.1 NANA data collection procedure 93](#_Toc513973739)

[4.2.2 Appropriateness of the data for predicting depression 96](#_Toc513973740)

[4.3 Descriptive exploration of the data 96](#_Toc513973741)

[4.3.1 Participant demographics 96](#_Toc513973742)

[4.3.2 Scores on the GDS 97](#_Toc513973743)

[4.3.3 Mood scores 98](#_Toc513973744)

[4.4 Methods 99](#_Toc513973745)

[4.4.1 Data cleaning 100](#_Toc513973746)

[4.4.2 Cluster analysis 101](#_Toc513973747)

[4.4.3 Regression 103](#_Toc513973748)

[4.4.4 Logistic regression 105](#_Toc513973749)

[4.5.1 Ridge regression 107](#_Toc513973750)

[4.5.2 Selection of variables using the LASSO 107](#_Toc513973751)

[4.5.3 ROC curve analysis 109](#_Toc513973752)

[4.6 Results 110](#_Toc513973753)

[4.8 Discussion 116](#_Toc513973754)

[4.8.1 Strengths and limitations 120](#_Toc513973755)

[4.9 Conclusion 121](#_Toc513973756)

[Chapter 5 - Older adults’ perspectives on technology to report on mood and support mental health 124](#_Toc513973757)

[5.1 Introduction 124](#_Toc513973758)

[5.1.1 Background 124](#_Toc513973759)

[5.1.2 Aims 126](#_Toc513973760)

[5.2 Research Questions 127](#_Toc513973761)

[5.3 Approach 127](#_Toc513973762)

[5.4 Pilot study 129](#_Toc513973763)

[Reflections 130](#_Toc513973764)

[5.5 Main study 132](#_Toc513973765)

[5.5.1 Design 132](#_Toc513973766)

[5.5.2 Participants 133](#_Toc513973767)

[5.5.3 Inclusion criteria 133](#_Toc513973768)

[5.5.4 Ethics and payment 134](#_Toc513973769)

[5.5.5 Procedure 134](#_Toc513973770)

[5.5.6 Analysis 139](#_Toc513973771)

[5.6 Results 141](#_Toc513973772)

[5.6.1 Template analysis 143](#_Toc513973773)

[Usability 144](#_Toc513973774)

[Motivators 149](#_Toc513973775)

[Barriers 152](#_Toc513973776)

[Awareness of technology to support mental health 155](#_Toc513973777)

[5.7 Discussion 156](#_Toc513973778)

[5.7.1 Principal Results 156](#_Toc513973779)

[5.7.2 Contribution to knowledge 161](#_Toc513973780)

[5.7.3 Strengths and limitations 163](#_Toc513973781)

[5.7.4 Summary 164](#_Toc513973782)

[Chapter 6 – Collection of primary data to validate depression models. 166](#_Toc513973783)

[6.1 Introduction 166](#_Toc513973784)

[6.1.1 Rationale 166](#_Toc513973785)

[6.1.2 Research question 167](#_Toc513973786)

[6.1.3 Study design 167](#_Toc513973787)

[Additional measures 167](#_Toc513973788)

[6.2 Method 168](#_Toc513973789)

[6.2.1 Participants 168](#_Toc513973790)

[6.2.2 Sample size 169](#_Toc513973791)

[6.2.3 Recruitment 170](#_Toc513973792)

[6.2.4 Ethics 173](#_Toc513973793)

[6.2.5 Pilot study 173](#_Toc513973794)

[6.2.6 Procedure 174](#_Toc513973795)

[6.2.6 Data analysis 177](#_Toc513973796)

[6.3 Results 180](#_Toc513973797)

[6.3.1 Participants 180](#_Toc513973798)

[6.3.2 Results from the ROC analysis 187](#_Toc513973799)

[6.4 Discussion 190](#_Toc513973800)

[6.4.1 Limitations 191](#_Toc513973801)

[Chapter 7 – Developing a model for the prediction of future anxiety in older adults 195](#_Toc513973802)

[7.1 Introduction 195](#_Toc513973803)

[7.1.1 Research question 196](#_Toc513973804)

[7.2 Method 197](#_Toc513973805)

[7.2.1 Data collection 197](#_Toc513973806)

[7.2.2 Data analysis 197](#_Toc513973807)

[7.3 Results 198](#_Toc513973808)

[7.4 Discussion 199](#_Toc513973809)

[Chapter 8 - Exploring the views of healthcare staff on use of technology for the prediction of future depression and anxiety in older adults 202](#_Toc513973810)

[8.1 Introduction 202](#_Toc513973811)

[8.1.1 Rationale 202](#_Toc513973812)

[8.1.2 Research Questions 204](#_Toc513973813)

[8.1.3 Approach 204](#_Toc513973814)

[8.2 Methods 205](#_Toc513973815)

[8.2.1 Setting 205](#_Toc513973816)

[8.2.2 Participants 206](#_Toc513973817)

[8.2.3 Ethics 206](#_Toc513973818)

[8.2.4 Procedure 207](#_Toc513973819)

[8.2.5 Data analysis 208](#_Toc513973820)

[8.3 Results 210](#_Toc513973821)

[8.3.1 Participant information 210](#_Toc513973822)

[8.3.2 Thematic Analysis 211](#_Toc513973823)

[8.4 Discussion 225](#_Toc513973824)

[8.4.1 Principal results 225](#_Toc513973825)

[8.4.2 Comparison with prior work 227](#_Toc513973826)

[8.4.3 Reflexivity 229](#_Toc513973827)

[8.4.4 Strengths and limitations 230](#_Toc513973828)

[8.4.5 Conclusions 233](#_Toc513973829)

[Chapter 9 – Discussion 236](#_Toc513973830)

[9.1 Introduction 236](#_Toc513973831)

[9.2 Summaries of the studies 236](#_Toc513973832)

[Study 1 236](#_Toc513973833)

[Study 2 238](#_Toc513973834)

[Study 3 241](#_Toc513973835)

[Study 4 242](#_Toc513973836)

[Study 5 244](#_Toc513973837)

[9.3 Overall summary 246](#_Toc513973838)

[9.4 Strengths and Limitations 248](#_Toc513973839)

[9.5 Reflections on the use of the medical model 257](#_Toc513973840)

[9.6 Implications of the research 259](#_Toc513973841)

[Implications for policy 259](#_Toc513973842)

[Implications for healthcare practice 261](#_Toc513973843)

[Implications for research 262](#_Toc513973844)

[9.7 Future directions 263](#_Toc513973845)

[9.8 Contribution to knowledge 267](#_Toc513973846)

[9.9 Conclusions 268](#_Toc513973847)

[References 271](#_Toc513973848)

[Appendices 285](#_Toc513973849)

[Appendix 1. Search terms for the first mapping review (Chapter 3A) 285](#_Toc513973850)

[Appendix 2. Search terms for the second mapping review (Chapter 3B) 288](#_Toc513973851)

[Appendix 3. Activity sorting cards used in Study 2 (Chapter 5) 292](#_Toc513973852)

[Appendix 4. Vignette cards used in Study 2 (Chapter 5) 293](#_Toc513973853)

[Appendix 5. ScHARR Research Ethics Committee approval letter for Study 2 294](#_Toc513973854)

[Appendix 6. ScHARR Research Ethics Committee approval letter for Study 3 295](#_Toc513973855)

[Appendix 7. Health Research Authority approval letter for Study 5 (Chapter 8) 296](#_Toc513973856)

[Appendix 8. Interview Schedule for Study 5 (Chapter 8) 297](#_Toc513973857)

# **Table of Figures**

Figure 1. Diagram showing the different studies within the thesis and how these are linked 24

Figure 2. Diagram showing the research questions for the review, situated between the main themes of thesis. 30

Figure 3. Flow chart showing process used to conduct the systematic mapping review. 51

Figure 4. PRISMA 2009 diagram showing selection of articles in the first mapping review 56

Figure 5. Bar chart showing number of studies found exploring the prediction of future occurrence of these mental health conditions. 67

Figure 6. Bar chart summarising design of studies included in the review. 69

Figure 7. PRISMA 2009 diagram showing selection of articles in Part B of the mapping review 78

Figure 8. Diagram showing potential application of a machine learning approach 92

Figure 9. Older adult self-reporting mood using the NANA Homesystem software on an Asus Eeetop 15 inch touchscreen PC 93

Figure 10. Study design for the NANA validation study. Adapted from (Astell et al., 2014). 94

Figure 11. Boxplot showing scores on the GDS at baseline (T1) and at follow-up (T2) in the NANA validation trial (Astell et al. 2014) 98

Figure 12. Variation in participants' scores given for different mood words across all days in week 1 mood reports. 99

Figure 13. Scatter graph showing an example of clustering. 102

Figure 14. Scatter graph with trend line. Demonstrates linear regression within a two-dimensional space. 104

Figure 15. Deviance plot showing deviance at different values of lambda for the model with a cut-off of five points on the GDS. 111

Figure 16. Coefficient plot showing contribution of each coefficient to the logistic model. 112

Figure 17. Receiver Operating Characteristic curve for predictive model derived using LASSO for the model with a cut-off of five points on the GDS. 113

Figure 18. Graph showing deviance plot for model using a cut-off of three points on the GDS. 114

Figure 19. Coefficient plot for model derived from using LASSO with the NANA validation data with a cut-off of three points on the GDS. 115

Figure 20. Graph showing ROC curve analysis for model derived from the NANA validation data with a cut-off of three points on the GDS. 116

Figure 21. Diagram showing the coefficients included in the two models developed using the LASSO. 119

Figure 22. Boxplot showing distribution of mood scores across all days and all participants for Study 3. 184

Figure 23. Boxplot showing spread of GDS scores for Study 3 participants at baseline and at follow-up 185

Figure 24. Bar chart showing type of device used to complete daily mood reports 187

Figure 25. Graph showing ROC curve for the model at a cut-off of five 188

Figure 26. Graph showing ROC curve analysis for model using a cut-off of three points on the GDS. 189

Figure 27. Coefficient plot for application of the LASSO to self-reported mood data and HADS anxiety scores 198

# **Table of Tables**

Table 1. Framework of factors affecting older adults' adoption of technology. From (Lee & Coughlin, 2015). 44

Table 2. Main search terms used across databases in the systematic mapping review 53

Table 3. Extractions from the systematic mapping review (Part A). 60

Table 4. Machine learning techniques used in papers included in the review (A). 66

Table 5. Types of input data used in studies included in the review (Part A). 68

Table 6. Search terms used in the second mapping review (Part B) 74

Table 7. Extractions from the systematic mapping review (Part B) 80

Table 8. Mood words shown to participants on the touchscreen in the NANA validation trial within the question "How \_\_\_\_ do you feel right now?".. 95

Table 9. GDS summary statistics from the NANA validation trial 97

Table 10. Frequencies of participants in the NANA validation trial valued as depressed according to the GDS at cut-offs of three and five points 97

Table 11. Table showing AUCs for models produced in other approaches to predicting future depression 118

Table 12. Activities used in the pilot study (Study 2). 130

Table 13. Activities undertaken in the main study (Study 2). 136

Table 14. *A priori* themes chosen before beginning the coding process (Study 2). 140

Table 15. Participant demographics.(Study 2) 142

Table 16. Final template including themes and subthemes after coding (Study 2) 143

Table 17. Participant characteristics table for the final participants (after data cleaning), compared to NANA (Study 1) participants.. 183

Table 18. Frequencies of participants in Study 3 scoring above the cut-offs of three and five on the GDS (15-item) at baseline and follow-up 186

Table 19. Job roles and experience of participants in Study 4 210

# **Abstract**

**Background**: Many older adults fail to seek treatment for depression and anxiety before crises occur due to stigma or misperception of the conditions as signs of ageing. The research presented in this thesis aimed to explore the potential application of machine learning and digital technology to predict future depression and anxiety status in older adults, in order to permit earlier intervention.

**Method**: A narrative review and two mapping reviews were conducted to provide context and characterise existing literature in the area. Secondary data consisting of 40 older adults’ self-reported mood and appetite scores were analysed with machine learning techniques to predict depression status according to the Geriatric Depression Scale (GDS) at nine weeks’ follow-up. Primary data were then collected from 38 older people to validate the models developed earlier, and to develop a new model for the prediction of anxiety. Interactive group sessions were held with 15 older adults to explore their views on the use of digital technology to support mental health. Interviews were conducted with community healthcare staff to consider implementation in healthcare.

**Results**: The predictive models had high predictive ability compared to studies presented in the mapping review. Using machine learning to predict anxiety status did not generate a useful model for this sample, suggesting the input data were not appropriate for this purpose. Interactive sessions with older adults raised important issues on usability and motivation. Interviews with healthcare staff permitted an exploration of existing practices and prior implementation of technology in a community healthcare setting.

**Conclusions**: The work in this thesis has contributed to knowledge by proposing a new method of predicting depression and anxiety in older adults, and has demonstrated the potential of this approach. This is supported by an exploration of older adults’ views of using digital technology to monitor mood and manage mental health, which has not been completed in prior work. The contribution of the work is further strengthened by the inclusion of a study which considers the implementation of the approach.

# Chapter 1 – Introduction

## Introduction

This chapter introduces the main topic areas explored within the thesis, and provides a rationale for the studies presented in later chapters. It also outlines the overall research question and aims for the research, before providing an overview of the thesis.

## 1.2 Rationale

### Depression and anxiety in older people

Depression and anxiety are common, debilitating conditions which affect people across all age ranges. These unpleasant conditions frequently co-occur (Nordhus, 2008), and negatively affect quality of life (Brenes, 2007; Lam et al., 2009). Estimates of the prevalence of depression in older adults vary. For Caucasians, the figure is between 0.9 and 49% (Djernes, 2006), dependent on the setting. Depression in older adults is problematic, since it is more common than dementia (Allan et al., 2014) and over half of cases of depression in older adults are first-onset cases (Fiske et al., 2009), meaning many are unaware of the condition that causes the unpleasant symptoms they experience. Anxiety disorders are also common in older adults. Again their prevalence is hard to evaluate, though it has been estimated at between 1.2% and 15% in community-based adults over age 60 (Bryant et al., 2008).

Left untreated, the consequences of depression and anxiety can be severe for sufferers and those close to them. Depression and anxiety are both risk factors for suicide (Conwell et al., 2002; O’Connell et al., 2004), and among those with depression, older adults are more likely than younger adults to complete suicide attempts (Rodda et al., 2011). Older adults with depression or anxiety are also known to be stigmatised by other older adults (Webb et al., 2009), and to be at increased risk for impaired cognitive functioning (Rodda et al., 2011; Yochim et al., 2013).

Although treatments for depression and anxiety are available and are known to be effective in older adults, many do not get the help they need. This is in part owing to a lack of help-seeking when older adults experience symptoms (Kessler et al., 2014). This can be because they misinterpret symptoms (Lee & Dugan, 2015), or are unaware they have a medical condition, putting the symptoms down to “getting old” (Conner et al., 2010), or because they fear the stigma attached to mental health conditions (Préville et al, 2015). There is thus a need for research to explore ways of improving existing methods for detection of depression and anxiety in this population.

### Early intervention

Early intervention in the progress of mental health conditions could help to reduce future crises. Anxiety has been associated with attempted suicide (Chartrand et al., 2012). Depression too is a predictor of suicide in older people, and a majority of those older people who commit suicide have depression (O’Connell et al., 2004). Depression and anxiety cause great emotional turmoil, thus early intervention is of benefit for the reduction of this turmoil (Halfin, 2007). The ability to predict future occurrence of these conditions is thus an important goal, particularly since research has shown that early intervention can reduce the length of episodes of depression (Kupfer et al., 1989).

The National Institute for Health and Care Excellence (NICE) offers evidence-based guidance on the management of depression and anxiety. It advocates a stepped care model, in which patients are differentially managed depending on the severity of their condition. Management at lower ‘steps’ may involve signposting to self-help resources or low intensity cognitive behavioural therapy (CBT), while patients at higher steps may require pharmacotherapy, CBT at higher intensity, or other forms of psychotherapy (NICE, 2016; NICE, 2011). Recommended treatments at lower steps are much less resource intensive than those at higher steps of the models and take less time to implement. Futhermore, research suggests that early intervention in depression improves prognosis (Halfin, 2007; Kupfer et al., 1989). Therefore early interventions in the course of depression or anxiety are likely to be less disruptive to the patient’s life, while improving prognosis and costing less to provide.

For some patients, pharmacotherapy is deemed necessary to encourage remission or management of depression or anxiety. However, different drugs affect people in different ways, and typically patients will have to try a number of different drugs before finding one that improves their condition with acceptable side effects. Intervening in these conditions early on can allow time for patients to try courses of different drugs before their condition worsens to crisis point.

### Current practices in the detection of depression and anxiety

Many clinicians use published scales to assist with the assessment of patients for anxiety and depression. These scales consist of questions that ask patients to think back over a recent time period (often the past two weeks) and consider how they have felt. The answers patients give are measured against scoring guides to establish whether they are experiencing depression or anxiety. Examples of these include the Patient Health Questionnaire 9 (PHQ-9) (Kroenke et al., 2001) for depression, and the Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983) for depression and anxiety. There also exist scales exclusively for older adults and these include the Geriatric Depression Scale (GDS) (Yesavage & Sheikh, 1983) and the Geriatric Anxiety Inventory (GAI) (Pachana et al., 2007). Since these are targeted at older adults, they omit questions focussing on symptoms which feature in other common conditions of old age. While the scales mentioned here have been validated in clinical populations, their reliance on patients’ own memory of how they have felt in the past reduces their reliability, as it is known that present mood state affects how past mood states are perceived (Barsky, 2002). It is important therefore to consider how these tests could be improved. Brevity and simplicity are likely to be important considerations in the development of new tools (Slade et al., 1999).

### Digital technologies and older adults

Modern digital technologies are growing in popularity among older adults in the UK. Ofcom, the independent regulator for the UK communications industry, reported in 2015 that the proportion of older adults (65+) who owned a smartphone had increased from 2% to 17% between 2010 and 2014, and over the same period, the proportion of adults aged 65 and over who used a tablet computer to go online increased from 1% to 17% (Ofcom, 2015). These figures suggest active take-up of new digital technologies among adults of retirement age in the UK. The report also suggests that smartphone and tablet usage is growing more rapidly in this age group than in any other.

Smartphone and tablet technologies are being employed increasingly in healthcare to enable contact between health professionals and patients, in a movement known as “telehealth” (Darkins & Cary, 2000). The concept of telehealth contains within it the concept of “telemonitoring”, which describes the use of technology for patients to record measures relating to their health and transmit these via a phone line, internet, or mobile network. This permits a reduced dependence on patients to visit a healthcare site, and allows data on a patient’s condition to be collected more frequently. As more and more older adults become comfortable with the idea of using digital technologies to communicate, there is increased potential for telemonitoring to be used for older adults to self-report on conditions relating to subjective mental states like anxiety and depression. This offers the possibility of overcoming some of the difficulties of one-time meetings to detect and diagnose these conditions, by allowing data to be reported over a number of days. This has the benefit of counteracting bias caused by present mood state (Barsky, 2002).

Many new applications of telemonitoring are being developed and evaluated by researchers. Applications where these have been successful include in diabetes, where it is found to have a beneficial effect on glycaemic control (Polisena et al., 2009), and chronic heart failure, where remote monitoring reduces rates of hospital admission by 21%, and reduces mortality by 20% (Clark et al., 2007). Telemonitoring has also been applied in the management of chronic obstructive pulmonary disease and has been found to reduce rates of hospitalisation in patients who use it (Polisena et al., 2010).

While telemonitoring can be used to monitor aspects of physical health such as blood pressure and blood glucose, it also has the potential to be used to monitor mental health, where subjective mood states can be recorded using self-report approaches. One example of a system that takes this approach is the NANA (Novel Assessment of Nutrition in Aging) homesystem. NANA is a validated, touchscreen system that allows older adults to report on their mood, appetite, exercise and food intake, among other measures (Astell et al., 2014). It was developed with an older adult user group and employs a user-friendly interface to allow new possibilities for remote data collection. This tool has been successfully used to collect data on six aspects of mood and appetite from 40 older adults over three, one-week periods (Astell et al., 2014). The data that can be collected using the NANA system may hold promise for the prediction of a number of health conditions, including depression and anxiety.

### Machine learning

Alpaydin (2004) states that “the goal of machine learning is to program computers to use example data or past experience to solve a given problem”. Machine learning has long been applied in the field of systems engineering, and more recently has been applied in the field of online retail, for example to provide recommendations to shoppers (Linden et al., 2003).

In recent years, researchers have begun to apply machine learning tools to solve problems in the field of healthcare. This includes applications in genetics and genomics (Libbrecht & Nobel, 2015) and in medical imaging (Wernick et al., 2010). Approaches have also begun to explore the detection and diagnosis of diseases, including mental health conditions (e.g. Karstoft et al., 2015; Passos et al., 2016; Jiménez-Serrano et al., 2015). The use of computers to process large datasets is helpful for finding patterns in past data to predict future outcomes. Applications of this sort hold promise for enabling the prediction of, and therefore early intervention in, mental health conditions, potentially preventing suicide or other adverse events. The increasing, widespread, voluntary use of digital technologies, including among older adults (Ofcom, 2015), facilitates the collection of data that could be used in this manner.

### Implementation of digital technologies within healthcare

For all the benefits that technology can bring, the introduction of new technologies in health and social care organisations can be problematic. While an increasing number of evidence-based apps and websites to support mental health are becoming available, research on their uptake into healthcare is limited (Drozd et al., 2016). Actual uptake tends to take many years to occur, with few research-based solutions reaching implementation (Mohr et al., 2017a). Some research demonstrates that implementation can fail when staff are not included in decisions about whether and how to take on new technological solutions (Kyratsis et al., 2012). Furthermore, when frontline staff do not have buy-in to new technological approaches to healthcare, effective uptake can become more difficult (Ackerman et al., 2012). It is thus of great importance to work with healthcare staff when developing new technologies to facilitate consideration of staff’s needs.

### Summary

Depression and anxiety in older adults are underdiagnosed and often overlooked, despite their high prevalence and potential for harm. Given this potential for harm, early intervention in mental health problems in older adults is important, and may help to prevent crises from occurring, and reduce costs associated with health care for those experiencing mental health conditions. Machine learning is increasingly being applied in many areas of healthcare. The use of machine learning to predict depression and anxiety in older adults could benefit society by increasing opportunities to intervene early on in the progress of these conditions. Measures of mental health to be analysed using a machine learning approach could be collected using digital technologies, and the rate of uptake of these devices by older adults is increasing. Prior work suggests that engaging with end users, as well as with healthcare staff from multiple roles and levels, during the design process, can improve chances of successful implementation of new technologies in healthcare.

## 1.3 Aim and objectives

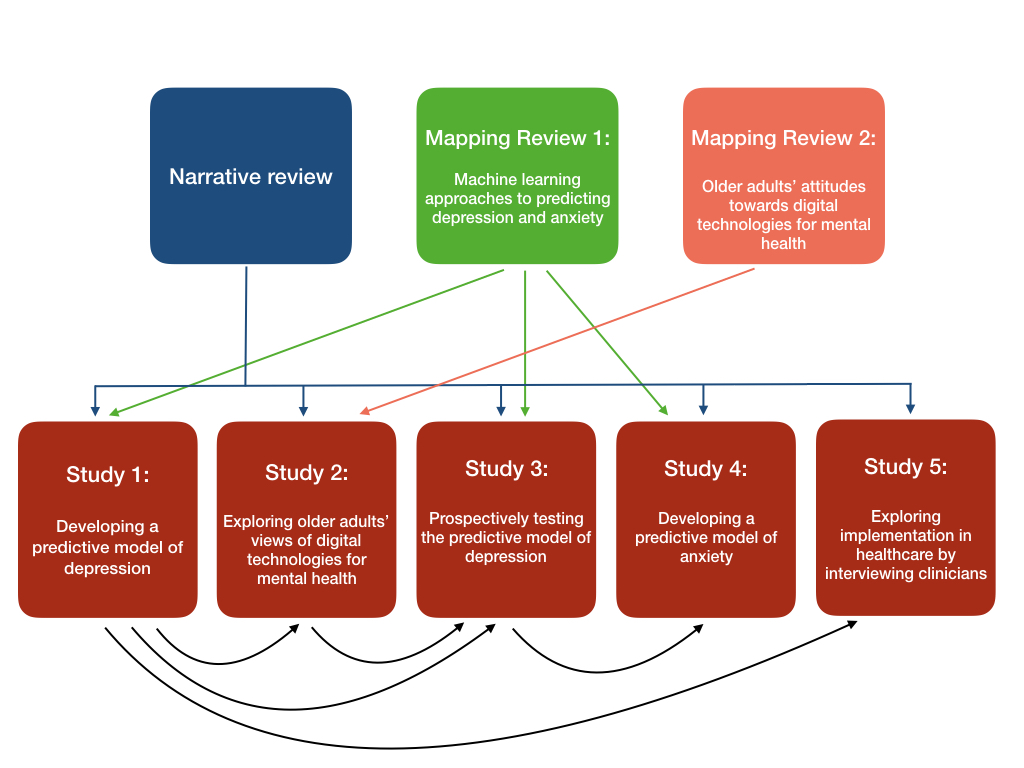
The overall aim for the thesis is to explore the practical application of machine learning to self-reported mood and appetite data, in order to predict future depression and anxiety in older adults, and explore the potential for using this approach in community healthcare.

The objectives of this thesis are as follows:

1. To explore the application of machine learning tools to predict depression and anxiety in older people, using secondary and primary datasets.
2. To explore older adults’ engagement with technology to monitor and support mental health.
3. To explore the views of healthcare staff on implementing technology to monitor older adults’ mental health and predict future depression and anxiety.

## 1.4 Summary of the research completed

The work in this thesis is comprised of literature reviews and empirical work to explore the application of digital technology to the prediction of depression and anxiety in older adults. Multiple methodologies were used to explore the area of interest, including literature reviews, quantitative and qualitative approaches. Figure 1 provides an overview of the studies completed for the thesis.



**Figure 1. Diagram showing the different studies within the thesis and how these are linked. Arrows indicate where findings from each study informed the development of methods for subsequent studies.**

The thesis begins with a narrative review of the literature (Chapter 2), which explores the themes of depression and anxiety in older adults, digital technologies to support mental health, older adults’ use of technology, early detection, and implementation. The literature presented in the narrative review informed the development of each of the empirical studies (1 to 5).

The first mapping review, presented in Chapter 3A, explores past approaches to the use of machine learning to predict depression and anxiety. A systematic approach was taken to characterise and categorise the literature in this area. The mapping review shows that a greater number of studies have been published that explore the prediction of depression than have been published on the prediction of anxiety. Reasons for this are explored in the chapter. In addition, the review found no studies that had explored the prediction of depression or anxiety in older adults, indicating a gap for research in this area. This review therefore demonstrates a need for the research conducted in studies 1, 3 and 4, which explore the possibility of predicting depression and anxiety in older adults.

The second mapping review, presented in Chapter 3B, systematically maps the literature on older adults’ attitudes towards the use of digital technologies to support mental health. The review demonstrates a lack of research in this area, and shows that the few existing studies in this area have a high level of bias. The findings from this review showed there was a research gap in literature for a study to explore older adults’ views of using digital technologies to support mental health. This informed the development of Study 2, which takes a qualitative approach to explore older adults’ views of using digital technologies to monitor mood and support mental health.

The first empirical study (Study 1) is presented in Chapter 4. This study explores an existing dataset that included self-reported mood and depression data for 40 older adults. The study involves the development of a machine learning approach to predict depression status from six measures of self-reported mood. Findings from the study indicated that a high level of predictive ability could be attained by applying particular machine learning techniques. This indicates promise for the approach. These results informed the development of the approaches taken in Studies 3 and 4. The finding that applying machine learning produced a model with good predictive ability also meant that it was worthwhile exploring how this approach could be implemented in technology with older adults and healthcare staff, thus supporting the need to conduct Studies 2 and 5.

Involving end-users in the development of new technologies is essential to understand their usability requirements. Thus, Study 2 consisted of interactive sessions with older adults which explored the usability of different apps and websites to monitor mood and support mental health. Activities in these sessions included card sorting, testing different apps and websites, and reflecting on vignettes. Template analysis was used to analyse the results. These showed that some participants currently use technology to alleviate negative feelings through music or video games, and that self-reliance was a motivator to engage with technology to support mental health. Participants had fears around what might happen when an alert was triggered. They described portability, screen size and clear instructions as important factors of usability. These findings were used to develop the data collection techniques used in Study 3.

The third empirical study in this thesis involved collecting data on mood and depression from a new sample of older adults online in order to validate the models developed in Study 1. Data collection methods were informed by older adults’ comments in Study 2. Two models were applied to newly collected data, with analyses conducted to evaluate their predictive ability. While the results showed a greater than chance level of prediction for both models, the low rate of depression in the new sample meant that the predictive analyses were low in quality. Further data collection from a larger sample was therefore recommended for future work.

Study 4 involved using the LASSO machine learning technique with the mood data collected in Study 3, along with data on participants’ later anxiety status, to develop a model for the prediction of future anxiety in older adults. This attempt was not successful, however, as the LASSO selected all mood variables out of the model, indicating that these were not predictive of future anxiety status. Further work could explore the use of different mood words more related to the experience of anxiety.

The final empirical study in the thesis, Study 5, is a qualitative exploration of the views of community healthcare staff on the implementation of digital technology for the prediction of depression and anxiety in older adults. Since the models developed in Study 1 and tested in Study 3 showed promise (above chance-level prediction of future depression), it was deemed important to explore the potential for implementing the approach in healthcare pathways. Seven healthcare practitioners were interviewed, and transcripts were analysed using Thematic Analysis. Findings included that pathways exist in community healthcare under the NHS in Sheffield within which a predictive approach could be implemented, and that the members of staff interviewed had experience using telehealth devices. While staff recognised benefits to the introduction of predictive technology for depression and anxiety, they also described a number of barriers, including the potential for staff to avoid engaging in new practices which might increase future workload. These mixed results were informative for approaches moving forwards.

The thesis ends with a discussion chapter, which reviews the work in the context of prior research, highlights the contribution to knowledge of the work presented, explores the strengths and weaknesses of the work, and makes recommendations for future research.

# Chapter 2 – Narrative literature review

## 2.1 Introduction

This chapter explores the key literature relating to the four main themes of this thesis: technology, older adults, depression and anxiety, and detection. A narrative literature review was conducted to understand the potential of using new technology to improve the detection of depression and anxiety in older adults. A further aim of the review was to highlight important work in related areas and to situate the thesis within a wider research context. The remit of the review is presented diagrammatically in Figure 2. This figure presents the research questions for the review at the intersections between themes.

How prevalent are depression and anxiety in older adults?

What treatments are available?

What are the correlates and risk factors for depression and anxiety in older adults?

How are depression and anxiety detected and diagnosed in older adults?

What are the barriers to detection?

What is known about older adults’ attitudes towards technology, including for healthcare purposes?

How is technology currently being used to detect and manage mental health conditions?

Figure 2. Diagram showing the research questions for the review, situated between the main themes of thesis.

## 2.2 Method

### Approach

The approach taken for this review was that of a narrative review (Coughlan, 2013; Cronin et al., 2008) including a narrative synthesis of findings (Booth et al., 2016). A systematic review was not deemed appropriate here because the aim was to establish a context within which to develop a research programme, rather than to exhaustively present and assess all research in a particular area. Accordingly, numbers of results were not recorded, and quality assessment of papers was not conducted.

### Process

A variety of methods were used to source the included literature. Many of the cited articles were found through searches of five research databases: Medline, PsycInfo, ASSIA, Web of Science and Google Scholar. Additional searches for articles, books and book chapters were conducted using the library catalogue of the University of Sheffield. These searches were supplemented by hand searching for books and book chapters in library buildings. Some books and papers were also found through recommendations from experts, and hand searching of reference lists also yielded useful results. Grey literature was obtained using Google searches, and through recommendations from experts. The search process was iterative, in that books and papers reviewed often provided insight and additional terminology which informed further searches on research databases and library collections. Papers, books and book chapters were included in the review where these were relevant to the research questions above. Systematic reviews were prioritised for inclusion wherever available and relevant. To reduce bias, papers with contradictory findings to those already found were purposefully included in the review. Findings were synthesised under headings related to each of the research questions. The review was initially conducted in 2015 as part of the confirmation review process. It was subsequently updated in 2017.

## 2.3 Depression and anxiety in older adults

### Depression and its prevalence in older adults

The 5th Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychological Association, 2013) uses the term ‘depressive disorders’ to refer to a number different conditions. The manual does not, however, define the term ‘depression’, despite its widespread use in the research literature. Meanwhile, the manual specifies that major depressive disorder is the ‘classic’ condition within the category of depressive disorders. As such, this thesis assumes that uses of the term ‘depression’ in the research literature refer to major depressive disorder where the term is not otherwise defined, and uses the term ‘depression’ to refer to major depressive disorder.

Major depressive disorder is characterised by low mood, loss of pleasure or interest in life, changes in appetite, sleep disturbances, change in speed of movements, feelings of tiredness, guilt and suicidal ideation. The DSM-5 suggests that five or more of these symptoms should be present over a two-week period for a diagnosis of major depression to be made (American Psychological Association, 2013).

The manifestation of depression can be different in older and younger adults. In older adults, depression tends to be more associated with increases in somatic symptoms, cognitive impairment, disability, utilisation of healthcare resources, suicide, and medical mortality (Rapoport et al., 2016).

Depression is common, estimated to affect 7% of the general population in a 12-month period (American Psychological Association, 2013). Studies examining the prevalence of depression in older adults vary widely in their findings [between 0.9 and 49%, (Djernes, 2006)], depending on the setting from which the sample is taken. However, it is generally considered that the prevalence of depression in older adults is lower than that in the general population (American Psychological Association, 2013). Nevertheless, depression is more common than dementia in older adults (Allan et al., 2014) and incidence of mild depressive symptoms in this population is found to be high (Büchtemann et al., 2012). Depression can be a recurrent condition, with a number of episodes occurring across the lifespan. It can be experienced for the first time at any age. Where this occurs in people over the age of 60, it is termed ‘late-onset depression’ (Driscoll et al., 2005).

### Anxiety and its prevalence in older adults

Anxiety disorders are also common in older adults, though their prevalence is hard to estimate. A review found estimates of between 1.2% and 15% of community-based adults over age 60 to be affected by anxiety (Bryant et al., 2008). Symptoms of anxiety include persistent, excessive worry which is difficult to control, restlessness, tiredness, problems with concentration, irritability, sleep disturbance and muscle tension (American Psychological Association, 2013). Anxiety commonly co-occurs with depression (Nordhus, 2008) and these conditions have a number of symptoms which overlap (Zbozinek et al., 2012). In addition, anxiety disorders are correlated with an increased rate of suicidal and self-injurious behaviour (Chartrand et al., 2012).

### Treatments for depression and anxiety

Treatments for anxiety and depression include pharmacological, cognitive and behavioural approaches, which can be delivered in a number of ways (Comer, 2013). The National Institute for Health and Care Excellence publishes guidance on the treatment and management of depression and anxiety (NICE, 2016; 2011). Recommendations are grouped by severity, with suggested treatment ranging from education and active monitoring to complex drug and psychological therapies.

The most commonly prescribed pharmaceutical treatments for depression for patients of all ages are selective serotonin re-uptake inhibitors (SSRIs) (American Psychological Association, 2013), which work on neurotransmitter receptors in the brain. Pharmaceutical treatments for anxiety also exist, although antidepressants are also commonly prescribed (Martín-Merino et al., 2010), as these also reduce symptoms of anxiety while having a lower risk of side effects (Comer, 2013).

The most common form of psychological therapy for anxiety and depression is cognitive behavioural therapy (CBT), an approach that works to change the thought processes and behaviours characteristic of depression and anxiety (Bhar & Brown, 2012). In recent years, online therapies have increased in number and there is now a movement of computerised CBT (cCBT) which involves using apps and websites to reflect on feelings and share experiences with other users. However, while some research supports the efficacy of online approaches (Andrews et al., 2010), some research has described a lack of evidence to support them (So et al., 2013). Despite the controversy around cCBT, NICE now includes cCBT as an option in its guidelines for the management of depression and anxiety (NICE, 2013).

### Risk factors and correlates of depression and anxiety in later life

There are many risk factors for the development of depression and anxiety, and awareness of these can help to effectively direct efforts to improve detection. These include the onset of physical health problems and stressful changes in personal circumstances. For example, it has been suggested that heart failure patients would benefit from screening for anxiety (Easton et al., 2015) and type 2 diabetes carries an increased risk of depression of nearly twice that of the general population (Roy & Lloyd, 2012). Additionally, both depression and anxiety are found to be common in chronic obstructive pulmonary disorder (COPD) (Yohannes et al., 2000). Stressful life events also increase risk, for example divorce is a recognised risk factor for poor mental health at all ages (Menaghan & Lieberman, 1986). While a certain level of distress is expected in response to major illness and difficult life events, the prolonged experience of low mood, reduced interest and low motivation after a reasonable period of time can be problematic.

The presence of mental health conditions may cause older adults to become greater consumers of physical healthcare. Unützer et al (2009) found that the presence of mental health problems in older adults correlated with increased healthcare costs, irrespective of mental health-specific treatment costs. Adams et al (2015) found that older adults with mental health problems spent longer in hospital per stay, had higher rates of readmission and spent longer waiting for emergency room treatment. Recognition of mental health problems in older adults at the earliest possible moment therefore may represent a way to reduce burden on all healthcare services, including, but not limited to, mental health services.

## 2.4 Detection and diagnosis of depression and anxiety in older adults

### Symptom identification and diagnosis

Depression and anxiety are formally diagnosed by a clinician in a process known as a structured clinical interview (First, 2016). However, symptoms of these conditions can be detected with questionnaire-type tests, and these can be used to aid diagnosis. Examples of tests for depression in current use include the Patient Health Questionnaire 9 (PHQ-9) (Kroenke et al., 2001), the Beck Depression Inventory (BDI) (Beck et al., 1961) the Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983), and the Geriatric Depression Scale (GDS) (Yesavage & Sheikh, 1983), the latter being designed specifically for the identification of depression symptomatology in older people. Questionnaire-type tests for anxiety include the General Anxiety Disorder 7 (GAD7) (Spitzer et al., 2006) and the Geriatric Anxiety Inventory (GAI) (Pachana et al., 2007), the latter being specifically for use with older adults. Some of these tools exist in digital form, for example Yesavage and colleagues have made an online version of the GDS (<http://www.medafile.com/GDS15.htm)>. The number of questions in each of these tests varies considerably, meaning the time needed to complete them also varies. Some tests have short versions which have been validated. For example, the original GDS consists of 30 questions, although a 15-item version of the test has also been developed and validated (Sheikh & Yesavage, 1986; Lesher & Berryhill 1994). Short tests can be beneficial because their administration time can be substantially lower than that of full scales (Almeida & Almeida, 1999).

### Barriers to detection and diagnosis

Detection of depression is known to improve prognosis, and failure to detect cases can increase mortality risk (Damián et al., 2017), thus it is important that barriers to detection and diagnosis are recognised. Primary care physicians may not pick up on the symptoms of poor mental health in older adults, since symptoms described by older patients are likely to include physical rather than psychological symptoms (Manthorpe & Iliffe, 2006). Moreover, symptoms of depression are often not recognised by older adults themselves as indicative of illness, particularly where depression is encountered for the first time in later life - older people often assume symptoms are merely part of the experience of ‘getting old’ (Smyer & Qualls, 1999), and therefore often do not seek help for the symptoms they experience (Crabb & Hunsley, 2006).

Stigma surrounding mental health conditions can also have a negative impact on older adults’ willingness to seek help. Stigma against individuals with mental health conditions can affect people at any age, and can occur in many different types of relationship and settings (Lasalvia et al., 2013). In older adults, this can take the form of discrimination, stereotyping or prejudice (Conner et al., 2010; Scazufca et al., 2016). Older adults tend to believe that fellow older adults who experience anxiety are responsible for the condition (Webb et al., 2009) and this stigma can negatively affect help-seeking behaviours, meaning that some older adults who experience these conditions do not receive treatment that may benefit them (Préville et al., 2015).

Fear of dementia is also known to affect help-seeking behaviours. Some symptoms of dementia overlap with those of depression in later life (Poison & Cory, 2013; Yesavage et al., 1982). Thus older adults experiencing symptoms of depression are known to avoid seeking help for symptoms of depression for fear of being diagnosed with dementia (Phillipson et al., 2015). This problem is particularly acute in older adults, since younger adults are less likely to develop dementia. Given the multiple and complex barriers to detection and help-seeking among older adults with depression and anxiety, it is timely to look at ways to improve the detection of these conditions.

## 2.5 Mental health technology

The development of smartphone and tablet applications and web technologies is allowing new approaches to be taken to the management and treatment of mental health conditions. Luxton et al (2011) highlight how smartphones can be used for patients to: conduct their own self-assessments; work with a therapist remotely via videoconferencing; access clinical knowledge; make use of a virtual coach for positive behaviour change; or enjoy games that reward healthy behaviour. Many organisations are already exploring ways in which digital technologies can complement or potentially improve upon conventional treatments for mental health conditions. These include Ieso Digital Health (iesohealth.com), which provides online therapy through a text chat system, and MindDistrict (Minddistrict.com), who provide videochat therapy supported via an app. From 2013-2015, the NHS supported a library of health apps which included a selection of apps to support mental health. Examples of these, highlighted in a report by Cotton et al (2014), included the Buddy app, a system of text messaging allowing users to keep track of how they feel and allowing doctors to access these data. They also included Wellhappy, an app facilitating access to NHS mental health services, and Five Ways to Wellbeing, an app encouraging reflection on daily practices to support mental wellbeing. Social media has also been used to support mental wellbeing, with several social media sites offering ways to monitor and share mood records with friends and relatives (Cotton et al, 2014).

However, while many apps designed to address symptoms of mental health conditions or improve coping mechanisms are available for download or purchase, little reliable evidence exists to support their efficacy. In a systematic review of the literature on apps for mental health, Donker et al (2013) found that the volume of mental health apps with no evidence base was extremely high compared to the number of evidence-based apps. In terms of health apps more generally, Cotton et al (2014) found there to be around 100,000 health apps available in major app stores, with no real quality filter apart from user reviews. Furthermore, scientific research into the efficacy of apps to support mental health is often flawed. While six of the eight papers reviewed by Donker et al (2013) found apps to be effective (i.e. there were significant improvements in participants’ levels of depression, stress or substance abuse as a result of using the apps), the quality of the papers reviewed was low. The design of the studies left results open to bias, for example, two studies failed to account for natural remission of symptoms (Burns et al., 2011; Rizvi et al., 2011). In terms of adherence to use of these apps over time, many of the studies involved payment to participants, which may have had a falsely positive effect on continuation rates. Furthermore, only one of the studies featured any sort of long term follow-up (Watts et al., 2013), meaning little is known about the sustainability of results from apps purporting to support mental health. There is therefore a need for new approaches to monitoring and managing mental health on digital platforms to be supported by scientific evidence.

Technology holds potential for the detection of mental health problems in individuals who would not otherwise seek help. Many older adults are unwilling to discuss their mental health, or are unwilling to visit their GP when they experience symptoms of anxiety or depression (Manthorpe & Illiffe, 2006). This may be caused by stigma around these conditions (Préville et al., 2015). With the rise in technology usage among older adults (Ofcom, 2015), technology could prove a useful way to enable older adults to report on their mood, allowing early intervention when their reports appear to indicate impending difficulties. This early intervention could take the form of self-care advice provided through the digital technologies themselves, or it could take the form of intervention from healthcare services, who could be alerted to the user’s potential decline in mental health via these technologies. Research to understand the views of both older adults and healthcare staff is required to understand the practicality and implications of these possible approaches.

## 2.6 Older adults and technology

Older adults appear to like using technology where it empowers them to do things more quickly or open up new possibilities. Chen and Chan (2013) asked adults aged 55 to 85 living in the community in Hong Kong about their use of technology in focus groups. Their participants found technology most beneficial for the convenience and speediness it provided for day-to-day tasks. Mitzner et al’s 113 community-based participants aged 65-85 gave examples of technological convenience they liked in the health domain such as being able to order prescription refills by phone, measure blood pressure at home, monitor their weight, and research health conditions (2010). For home life they mentioned enjoying use of technology for communication, cooking and entertainment. Peek et al (2015) describe how their participants enjoyed an improved sense of safety through technology, for example by carrying a mobile phone.

Studies with older adults have also covered the subject of motivation to use technology. Researchers have elicited a number of reasons that older adults would not be motivated to use technology. These include a belief that using technology requires a great deal of physical and mental effort (Mitzner et al., 2010), and that older adults do not have time to learn to use new devices (Chen & Chan, 2013). Other reasons include that technology is expensive (Greenhalgh et al., 2013; Chen & Chan, 2013), that it is often clunky or takes up too much space (Greenhalgh et al., 2013; Peek et al., 2015), or that older adults just prefer the old-fashioned ways of doing things (Peek et al., 2015). For participants in these studies, technology also aroused a number of fears, including about how technology might affect health, for example through radiation, eye strain or headaches (Chen & Chan, 2013). Older adults also have fears about security in use of technology, for example when using a bank card at a cashpoint (Chen & Chan, 2013). Furthermore, they fear addiction to technology (Peek et al., 2015; Chen & Chan, 2013).

Many older adults invest in technologies that they later discard or store and no longer use. Factors affecting older adults’ continued use of technologies are reported to be: reliability of the technology (Mitzner et al., 2010); having problems remembering how to use different technologies (Chen and Chan, 2013; Peek et al., 2015); relatives and close contacts being too busy/impatient/fast-speaking to teach older adults how to use technology effectively (Greenhalgh et al., 2013; Chen & Chan, 2013); and a view that older adults are able to manage on their own, or would prefer to manage their daily lives with the support of family rather than rely on assistive technology (Peek et al., 2015).

Thus, while older adults do see benefits to the use of technology and enjoy using it for particular purposes, it may be difficult for them to overcome the challenges that technology presents. This seems to affect older adults’ motivation to engage with technology.

Lee and Coughlin’s review of studies on older adults’ engagement with technology (2015) describes a framework of ten factors affecting adoption of new technologies. These factors and their associated descriptions can be found in Table 1. This framework can be used to better understand why older adults may take to certain technologies but not to others.

Table 1. Framework of factors affecting older adults' adoption of technology. From (Lee & Coughlin, 2015).

|  |  |
| --- | --- |
| **Factor** | **Description** |
| Value | Perception of usefulness and potential benefit |
| Usability | Perception of user-friendliness and ease of learning |
| Affordability | Perception of potential cost savings |
| Accessibility | Knowledge of existence and availability in the market |
| Technical support | Availability and quality of professional assistance throughout use |
| Social support | Support from family, peers, and community |
| Emotion | Perception of emotional and psychological benefits |
| Independence | Perception of social visibility or how a technology makes them look to others |
| Experience | Relevance with their prior experiences and interactions |
| Confidence | Empowerment without anxiety or intimidation |

Many of the factors presented in this framework appear to reflect the feelings of older adults as discussed in the papers described above. For example, ‘value’ is linked to the convenience and speediness mentioned by participants in (Chen & Chan, 2013), while ‘usability’ covers comments in (Peek et al., 2015) that older adults find it hard to remember how to use technologies, and ‘affordability’ in terms of the expense of technology is mentioned by participants in (Greenhalgh et al., 2013) and (Chen & Chan, 2013).

It is unknown if these factors may be relevant to the specific area of older adults’ use of technology to support mental health - none of the studies reviewed by Lee and Coughlin discussed such use of technology by older adults. As mentioned in Chapter 1, mental health is a stigmatizing issue for many older adults (Webb et al., 2009; Scazufca et al., 2016), so there may be some difference in attitudes towards technology for these purposes and those for technology more generally. For those involved in the development of technology for these purposes, as well as for those commissioning new technologies in mental health care, there is a need for research to explore factors that affect older adults’ adoption of digital technologies for self-reporting mood and managing mental health.

## 2.7 Implementation of technology in healthcare

Much research has demonstrated that consideration must be given to the implementation of technologies for use within healthcare services during their development. Failure to do so can result in low uptake (De Weger et al., 2013) or subsequent abandonment of new technologies (Ackerman et al., 2012). The reasons for this are multi-faceted.

Some research has shown that implementation failures can occur when healthcare staff at all levels are not involved in the adoption process. Ackerman et al (2012) examined how a kiosk system developed in hospitals for the automatic diagnosis of urinary tract infection (UTI) in women had been adopted by a healthcare trust, but subsequently abandoned. It was found that nurses had a generally low opinion of the kiosks due to problems with their functionality, which were unrecognised by senior staff, and this discouraged the nurses from directing patients to use them, despite monetary incentives being provided to the nurses for every UTI diagnosis made using the kiosks. Ackerman et al suggest that if nurses in the hospital trust had been involved in the design of the kiosk system, there may have been a greater chance of successful continued implementation of the technology.

Kyratsis et al (2012) examined how provision of different types of information about new technology to key staff members at the adoption stage differentially affected its successful implementation. They found that where frontline staff were not provided with “how-to” knowledge at adoption, that is, knowledge regarding the day-to-day use of the technology, these technologies were frequently abandoned. However, they found that senior staff required ‘principles’ knowledge (knowledge about the evidence base and inner workings of a technology) in order to support the adoption of such technologies. Thus, attempts to implement new technologies should consider presenting information about the intervention differently to different stakeholders.

Clinical ‘buy-in’ in from staff is also known to be important for successful implementation. Taylor and colleagues (2014) interviewed 105 members of healthcare staff from front-line and management roles about the implementation of telehealth systems. Their findings showed that the process of achieving buy-in could be hindered by staff experiencing difficulties while using telehealth. However, first-hand experience of the benefits of these technologies helped to develop trust in the systems. The authors suggested that developers could design telehealth services to allow incremental learning of the systems. They suggested too that providing healthcare staff with guidance on using telehealth systems, including the impact it has on patients and on care more generally, may help to improve adoption of these technologies.

A number of recent studies have explored the implementation of technology specifically in the field of mental health care. Drozd et al (2016) conducted a scoping review on the implementation of internet interventions for depression. Their findings included that brief training for staff is insufficient to sustain changes in practice over time, but that regular supervision by experienced and qualified staff can facilitate behaviour change among staff. This suggests that comprehensive training and support from experienced staff is important for successful implementation.

Mohr et al (2017a) state that the development of new interventions often overlooks aspects of human support of technological systems. They discuss how the development of many tools does not involve input from end-users, thereby failing to consider how new interventions might fit into users’ lives. Their own model for the development of new technological interventions, titled ‘accelerated creation to sustainment (ACTS)’, advocates using codesign with stakeholders, as well as working with clinicians to evaluate new technologies in real-world settings. Keeping the goal of implementation in mind throughout the development process is, they argue, essential to make best use of research and development of new interventions.

Ramsey et al (2016) conducted a survey of the administrative leadership of community behavioural health care agencies in the United States, to identify barriers to their use of technologies for behavioural healthcare. Barriers were found to include cost, privacy, users’ attitudes towards technology, organisational budgets and provision of good quality internet access locally. Their paper highlights a need for rigorous cost analysis of new interventions, and also highlights the importance of research considering the added value a technological intervention can provide to a healthcare provider. The paper also highlighted a lack of knowledge on the part of practitioners regarding the availability of different technologies they could employ. This underlined the need for dissemination and education about technological interventions across healthcare organisations.

It is clear from these studies that consideration of implementation from the outset is important when developing new technological interventions. Development should include co-design with end users and healthcare staff, as well as careful analysis of cost effectiveness and added value. Successful implementation is likely to also require careful consideration of the processes that surround the use of technologies within existing healthcare pathways.

# Chapter 3 (A) – Systematic Mapping Review 1

## 3.1 Introduction

Chapter 1 presented a rationale for the exploration of machine learning as an approach to the prediction of future depression and anxiety in older adults. This chapter systematically examines the literature to explore what has already been done to this end, and to discuss how new empirical research might usefully add to the currently available body of knowledge.

### Aims

This literature review took the form of a systematic mapping review. The review sought to characterise the literature available within the remit of the search criteria, and to establish where research gaps lay. The aim of the review was to explore the type of machine learning tools that had previously been applied to data on depression and anxiety. The review did not seek to assess the quality of the articles retrieved, since the aim was only to define and categorise existing approaches to the areas of interest. Neither did it seek to compare the predictive ability of the models developed using these techniques, as measures of predictive ability vary widely within the research literature. As such, this was not a traditional systematic review, but rather a systematic mapping review (Booth et al., 2016). The flow-chart in Figure 3 depicts the main stages involved in the review.

Search terms generated

Research question developed

Search terms refined with help from experts

Data-bases searched and findings sifted

Relevant details extracted to provide map of the literature

Figure 3. Flow chart showing process used to conduct the systematic mapping review.

### Research Question

The question to be explored in this review was as follows:

* What approaches have been taken to use machine learning to identify early symptoms of depression and anxiety?

## 3.2 Methods

### Decision of search terms to include

Firstly, key words in the research question were selected. Then, relevant articles known to the researcher were used to generate further search terms in a process known as ‘pearl growing’ (Booth et al., 2016). Using words and phrases from these articles ensured that appropriate terminology was used in order to capture as many relevant articles as possible.

A preliminary search with an initial set of search terms was run in Medline. Terms like “anxiety disorder” and “mood disorder” brought up many results which were subtypes of anxiety or depressive disorders (e.g. PTSD, panic disorder, agoraphobia) which were not the target of this search. Furthermore, ‘predictive ability’ was used as an alternative term for ‘machine learning’, although this returned many results which were not machine learning-related. Therefore, these terms were left out of the main search.

Experts were then consulted to discuss search terms chosen, and to expand on these where appropriate. These experts were: a professor with extensive experience of applied machine learning; a professor of health services research; and a research associate specialising in mental health, all of whom were experienced in the production of systematic reviews.

These experts suggested that key, known articles should be found when searches were run. Thus, the terms ‘prediction’ and ‘predict’ were added to the final set of terms, as these terms returned previously known articles. This had the added effect of increasing the number of results returned. Experts also advised including the term ‘neural networks’ in the search.

### Final search terms

Below are presented the main keywords used in this search. While the keywords searched in each database were the same, search operator characters varied across different databases, as did the availability of thesaurus words. For reproducibility, the exact search terms that were used within each database can be found in Appendix 1, along with the dates that each search was carried out.

Table 2. Main search terms used across databases in the systematic mapping review

|  |
| --- |
| Machine learning or neural network(s) or random forest(s) or lasso |
| AND |
| Early sign(s) or early symptom(s) or prodrome or first symptom(s) or predict or prediction |
| AND |
| Depression or depressive disorder(s) or depressive episode(s) or anxiety or generali(s/z)ed anxiety disorder |

### Databases searched

The following databases were searched. These were chosen for their relevance to the fields of mental health and technology.

* Embase
* MEDLINE via Ovid
* PsycINFO via Ovid
* ASSIA
* Web of Science

Google Scholar was also explored as part of the search process. However, this search engine did not provide enough space in its command line for all of the keywords and Boolean operators to be included in the search. There was also no function to combine searches. Some research has highlighted the functional limitations of Google Scholar and questioned its suitability for reviews with systematic approaches (Boeker et al., 2013). Because of these factors, Google Scholar was excluded as a search database for this review.

### Inclusion/exclusion criteria

Studies were only included in the review if they met all of the following criteria:

* Journal article/full paper conference proceeding
* Written in English language (to reduce the time and expense associated with translating foreign language papers)
* Reports on using techniques with a machine learning approach
* Discusses prediction of future onset/future occurrence of these conditions.
* Data may come from participants of any age (the focus here is on techniques, rather than examining evidence particularly in one population).

Exclusion criteria were as follows:

* Not future prediction
* Not focussed on depression or anxiety
* Not human data
* Not an empirical study

### Search technique

The review used free text searching (Booth et al., 2016) to search the bibliographic databases listed. Databases were chosen based on relevance to the research question. Thesaurus functions and subject headings were applied where available in each database. See Appendix 1 for full search terms used in each database.

### Sifting

Citations and abstracts of articles found in the searches were collected in a folder using the Mendeley software package. These articles were then sifted in accordance with the inclusion criteria. This was firstly done by title, then by abstract, and finally by full article. Where titles and abstracts were ambiguous, papers were kept in for the following round of sifting (i.e. ambiguous titles were considered at abstract level, and so on). From the final set of included papers, a pre-determined set of details were extracted from each paper and typed into an excel spreadsheet in order to build an overview of the literature in the area.

## 3.3 Results

The combined search results for all databases gave a total of 1239 articles. 285 duplicates were removed, leaving 954 articles for sifting. A PRISMA diagram of the selection of papers is presented in Figure 4.

Records identified through database searching:

(n = 1,239)

Records screened:

(n = 954)

Full-text articles screened for eligibility:

(n = 36)

Studies included in the information extraction:

(n = 19)

Records after duplicates removed:

(n = 954)

Records excluded:

(n = 918)

Full-text articles excluded:

(n = 17)



Figure 4. PRISMA 2009 diagram showing selection of articles in the first mapping review

### Title sift

The first sift was using titles only. In this sift, 844 articles were removed, for the following principal reasons. Numbers in brackets indicate the number of articles excluded for each reason.

* The research did not relate to depression or anxiety – the paper either discussed a sub-condition (e.g. PTSD) or used a broader term (e.g. low mood) (448)
* The research did not involve prediction of future onset of either condition (178)
* The aim of the research was not to predict the onset of the condition, but rather to predict something else (e.g. treatment response) (111)
* The research was not conducted with human data (73)
* There was no use of machine learning (21)
* The article was not from an academic journal or full paper conference proceeding (13)

### Abstract sift

After sifting titles of the articles, 110 articles were left to sift at the abstract level. Seventy-four articles were removed at this stage for the following principle reasons:

* The article’s use of the term “prediction” involved using one variable to predict another collected simultaneously, rather than seeking to make predictions about future depression/anxiety status from data collected at an earlier date (37)
* The main focus of the article was not on depression or anxiety but on some other condition (14)
* The object of prediction was not the onset of depression/anxiety, but rather something else, (e.g. treatment response) (11)
* The research was not conducted with human data (4)
* The research did not take a machine learning approach (4)
* The paper was not a peer-reviewed article (2)
* The article did not present the results of an empirical study (2)

### Full paper sift

The abstract sift left 36 articles for sifting at full paper level. During this sift, 17 papers were removed for the following principle reasons:

* The article’s use of the term “prediction” involved using one variable to predict another collected simultaneously, rather than seeking to make predictions about future depression/anxiety status from data collected at an earlier date (4)
* The object of prediction was not the onset of depression/anxiety, but rather something else, (e.g. treatment response) (4)
* The main focus of the article was not on depression or anxiety but on some other condition (2)
* The article did not present the results of an empirical study (3)
* The article focussed on predicting aspects of emotion or depressive tendency rather than the conditions of depression or anxiety (2)
* The research did not take a machine learning approach (2)

After this final sift, 19 papers were left to include in the information extraction stage of the review.

### Information Extraction

This part of the review involved extracting key information from each of the included studies to characterise the approaches taken from a broad perspective. The following details were extracted from the full papers:

* Machine learning tool applied (e.g. random forests, cluster analysis)
* Condition predicted (depression/anxiety)
* Input variable(s)
* Output variable
* Study design (prospective/retrospective)
* Measure of predictive ability
* Notes

Each paper was read multiple times to ensure accurate extraction of information. Table 3 shows details for each of the included papers. Papers are presented alphabetically by first author.

Table 3. Extractions from the systematic mapping review (Part A).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Paper reference | Machine learning tool applied | Condition predicted | Population | Source of input variables | Source of output variable | Prospective /retrospective study design | Measure(s) of predictive ability |
| (Alam et al., 2016) | Hidden Markov Model, Viterbi path counting, Stochastic Variational Inference-based training algorithm | "Emergency psychiatric state" (severe depression and anxiety) | 55 mental healthcare patients (age 20-66) | Biosensor data from nodes fitted on participants' skin | BDI, Scale for suicidal ideation, Beck hopelessness scale, Stress and Coping self-test, Impulsiveness scale, Aggression questionnaire, Hostility, PHQ9, GAD7 | Retrospective | True and false positive rates, Diagnostic odds Ratio |
| (Aldarwish & Ahmad, 2017) | Support vector Machine, Naïve Bayes | Depression | Depressed/non-depressed social media posts (no demographic data or quantities provided) | Posts from LiveJournal, Twitter, Facebook. | Researcher-defined sentiment (depressed or not depressed) | Unclear | Accuracy/Precision/recall |
| (Andrews et al., 2017)\* | LASSO | Depression | 40 older adults, 7 of whom reported having depression (age 65+) | Self-reported mood scores for 6 variables | Geriatric Depression Scale - 15 item depression status (score over 5) | Retrospective | ROC curve, AUC |
| (Bian et al., 2017) | Covariance regularised linear discriminant analysis | Depression and anxiety | 1000+ electronic health records of college students from 10 US universities (demographic info not specified) | Electronic health records | Diagnosis of mental health condition | Retrospective | Accuracy/F1-score |
| (Brann et al., 2017) | Logistic regression, LASSO, elastic net | Postpartum depression | 291 pregnant women, 63 with pp depression symptoms, 228 without (mean age 32.6) | 92 inflamation associated markers, and a summary variable (in blood) | Edinburgh postnatal depression scale and/or mini internat. neuropsyc. interview | Retrospective | Statistical significance of difference between depressed individuals and controls. |
| (Dabek & Caban, 2015) | Neural network model with hidden layer and SVM layer | Psychological disorders, including anxiety, depression, PTSD and behavioural/emotional states | 89,840 people with traumatic brain injury (demographic info not specified). | "Encounters" - doctor's appointments where a particular condition was diagnosed | Diagnosis of psychological disorder | Retrospective | ROC curve, AUC, accuracy |
| (Devi & Kumar, 2016) | First order Sugeno fuzzy logic model (neural network) | Anxiety | 36 Postgraduate students in India (age unspecified) | Moudsley Personality inventory - measures of neuroticism and extraversion | Sinha's comprehensive anxiety test | Retrospective | Mean absolute percentage error, root mean square error |
| (Foland-Ross et al., 2015) | Support vector machines, with 10-fold cross-validation | First onset of major depression | 33 adolescent females (aged 10-15) with no previous disorder | Baseline cortical thickness | Kiddie Schedule for Affective disorders and Schizophrenia for School-Age Children - Present and Lifetime version | Retrospective | Binomial test to see if model performed above chance. |
| (Hu et al., 2015) | Logistic and linear regression | Depression | 10,100 Social media (Weibo) users publishing 2.84-40 average daily posts, with 532 microblogs since registering. (No demographic info specified) | Linguistic and behavioural features including no of words, types of punctuation, no of pronouns; freq. of emotive words. | Depression based on CES-D | Retrospective | Precision and correlation coefficient |
| (Jimenez-Serrano et al., 2015) | Naïve Bayes, Logistic Regression, Support vector machines, artificial neural networks (Naïve Bayes found to be most predictive) | Postpartum depression | 1397 Caucasian women in a Spanish government funded study of postpartum depression (mean age 32.13) | Socioeconomic variables, personal and family history of depression, Eysenck Personality Questionnaire (for a measure of neuroticism) | Total score on Spanish version of the Edinburgh postpartum depression scale at week 8 and week 32 after delivery - cut-off 9 or more, probable cases taken forward for clinical interview to confirm | Retrospective | AUC, sensitivity and specificity, geometric mean of the two accuracies |
| (Jin et al., 2015) | Logistic regression with ridge parameter, multi-layer perceptron, support vector machine, random forest | Comorbid depression in diabetes | 1793 patients with diabetes from two clinical trials, primarily Hispanic, balance of depressive and non-depressive individuals (no demographic info specified) | 7 from 20 candidate predictors were: female, diabetes self-care, no of diabetes complications, prev. diag. of major depression, no of ICD-9 diag. in last 6 months, chronic pain, self-rated health | PHQ-9, PHQ-2 | Retrospective | AUROC |
| (Kessler et al., 2016) | Ensemble regression trees and 10-fold cross-validated penalized regression | Persistence, chronicity and severity of Major Depression | 5001 respondents to two surveys, across a 10-12 year period. (Age 15-54) | Comorbid lifetime disorders, parental depression, MDD incident episode symptoms, info about incident episode (age at onset and triggered/endogenous nature of episode), suicide attempts | Persistence-severity, incidence of hospitalisation for depression after first episode, disability status (due to depression) | Prospective | AUC, sensitivity, positive predictive value, likelihood-ratio positive |
| (Long et al., 2014) | Multi-variate pattern analysis - support vector regression, with leave-one-out cross validation | Stress-related disorders - depressive or anxiety symptom progression | 44 survivors of the Wenchuan earthquake, and 44 healthy controls from a city far from the earthquake (Mean age 37) | Amplitude of low-frequency fluctuation of particular neurological signals, and functional connectivity | Self-rating Anxiety Scale / Self-Rating Depression Scale | Retrospective | Permutation test |
| (Mourao-Miranda et al., 2012) | Gaussian process classifiers | Anxiety, depression | 32 adolescents (mean age 15), 16 who had a parent diagnosed with bipolar, depression or an anxiety disorder, 16 whose parents did not | Spatial patterns of brain activation during task | Result of diagnostic interviews with adolescents and their parents regarding their depression/anxiety symptoms | Prospective | ROC curve, permutation test |
| (Ooi et al., 2013) | Gaussian mixture modelling, Bayesian classification | Depression | 245 adolescents (age 12-13) | Glottal, prosodic, energy and spectral features of voice recordings from later-depressed and later-non-depressed participants | Depression status as rated by a psychologist using "conventional diagnostic tests" | Retrospective | Sensitivity, specificity and accuracy |
| (Rude et al., 2010) | Binary logistic regression | Major Depressive disorder | 44 Women in the US (age 25-63) | Dysfunctional attitudes scale score and scrambled sentences test score | Diagnosis of depression (Beck Depression Inventory and DSM Clinical Interview) in the follow-up period | Prospective | Significance and R squared of model prediction |
| (Schalinski et al., 2016) | Conditioned random forest regression | Depression (+ PTSD and dissociative symptoms) | 129 adult inpatients at a psychiatric hospital (mean age 26.1) | MACE scale of childhood adversities at different ages | Hamilton Depression Scale, based on observations of patients, patient self-reported symptoms and healthcare assistant's opinions | Retrospective | Akaike Information Criterion |
| (Tortajada et al., 2009) | Multi-layer perceptron (Artificial neural network) with weight pruning and holdout evaluation | Postpartum depression at 8 and 32 weeks after delivery | 1397 Caucasian women in a Spanish government funded study of postpartum depression | Psychiatric and genetic information (re: serotonin transporter gene), sociodemographic variables, EPDS just after childbirth, neuroticism on Eysenck Personality Questionnaire short scale | Total score on Spanish version of the Edinburgh postpartum depression scale at week 8 and week 32 after delivery - cut-off 9 or more, probable cases taken forward for clinical interview to confirm | Prospective | AUC and geometric mean |
| (Wardenaar et al., 2014) | Regression tree analysis, LASSO, ridge and elastic net penalized regression, k-means cluster analysis | Persistence (10 yrs+) and chronicity (15 yrs+) of Major Depression, hospitalisation and disability | 8261 respondents to the WHO world mental health surveys, with lifetime major depressive disorder (no demographic information available) | Incidence of other mental health conditions according to the survey | Measures relating to duration of major depressive condition and hospitalisation and disability | Retrospective | Sensitivity, positive predictive value, likelihood ratio positive |

**\*This article was written by the researcher and his supervisory team, and discusses research reported in a later chapter of this thesis. Therefore, it has been excluded from the analysis and discussion of the literature review below.**

### Characterisation of studies included

A wide variety of machine learning techniques have been applied to the prediction of depression and anxiety. Table 4 shows the techniques used in the papers included in the review. The most commonly used techniques among the studies were support vector machines and logistic regression, while neural networks, Bayesian methods, random forests, elastic net and LASSO were also each used in more than one study.

Table 4. Machine learning techniques used in papers included in the review.

|  |  |
| --- | --- |
| **Technique used** | **Number of studies** |
| Support vector machines | 5 |
| Logistic regression | 5 |
| Neural networks | 4 |
| Bayesian methods | 3 |
| Random forests | 2 |
| Elastic net | 2 |
| LASSO | 2 |
| Other (*Techniques with only 1 incidence of use among papers in the review*) | 7 |

The type of condition predicted also varied among the papers. Figure 5 shows a graph showing the number of studies found with the aim of predicting depression, anxiety, and multiple conditions. The majority of the studies included in the analysis (12 out of 18) focussed on predicting future depression only. Just one study focussed on predicting future anxiety only. Five papers reported on approaches to predict the occurrence of multiple conditions, including both depression and anxiety.

Figure 5. Bar chart showing number of studies found exploring the prediction of future occurrence of these mental health conditions.

Study samples varied widely in age and sample size. Sample size varied from a minimum of 32 to a maximum of 89,840. Eight of the 18 studies had sample sizes of over 1,000. Many studies failed to report demographic information about their participants. None of the studies reported recruitment from an exclusively older adult population.

Table 5 shows the types of input data that were used to train the models used in the studies. The most common source of input data for the machine learning models was records of medical histories. A small number of studies used one-time, self-report scales or neurological measures such as MRI data as the input data for their analysis. Less commonly used data sources included social media posts, voice recordings and biosensor data.

Table 5. Types of input data used in studies included in the review.

|  |  |
| --- | --- |
| **Type of input data** | **Number of studies** |
| Medical history | 6 |
| One-time self-report scales | 3 |
| Neurological measures | 3 |
| Social media posts | 2 |
| Voice data | 1 |
| Blood markers | 1 |
| Biosensor data | 1 |
| Multiple | 1 |

Output variables were mostly self-report measures of depression or anxiety, although six of the 18 papers in the analysis used a clinical diagnosis of one of the conditions as the output variable in their analysis. Techniques used to measure predictive ability included receiver operating characteristic (ROC) curve analysis, sensitivity, specificity, precision, recall and accuracy. Figure 6 shows numbers of prospective and retrospective studies included in the analysis. The majority of studies included in the analysis (14) were retrospective in design, while only four had a prospective design.

Figure 6. Bar chart summarising design of studies included in the review.

## 3.4 Discussion

This review took a systematic mapping approach to explore existing literature on the use of machine learning methods to predict future depression and anxiety. The aim was to establish the types of technique which are currently in use, and to understand where there were research gaps in the existing body of research. After sifting by abstract, title and full paper, 19 papers were included in the review. This included one paper written by the author of this thesis, which was excluded from the analysis since it referred to work reported in later chapters of this document.

Support vector machines, logistic regression and neural networks were the most commonly used techniques in the papers included in this review. The high number of studies reported to be using neural networks may relate to the inclusion of neural networks as an explicit search term in the search strategy. Bayesian methods, random forests, elastic net, and LASSO were each used in more than one of the papers. As such, a wide variety of machine learning tools have been applied to the prediction of future depression and anxiety.

This review found no study attempting to apply machine learning techniques to predict future occurrence of depression and anxiety in an exclusively older adult population. While some studies reported that their samples included one or more people over the age of 65, some studies failed to report demographic information entirely, and others reported the age range of their sample being lower than 65. The fact that no study had an exclusively older adult sample is important, as mood trajectories are known to vary between older adults and younger adults (Stanley & Isaacowitz, 2011). The lack of interest in the over-65 age group may reflect the fact that these conditions are reported to be less common among older adults than among younger adults (American Psychological Association, 2013). However, given that depression is more common than dementia among older people (Allan et al., 2014), and given too that both depression and anxiety are correlates of suicide (Chartrand et al., 2012; Paplos et al., 2003), being able to predict future occurrence of these conditions in this population is an important goal.

The review shows too that many approaches to prediction have focussed on identifying high risk by exploring static, and often short-term measures taken over a period of less than one day. The use of, for example, neurological measures, blood markers and one-time self-report scales as input data in a machine learning process may not be as reliable as approaches taken over multiple days because they rely on data from a single point in time. This means that data could be affected by short term events which affect the body and mind temporarily but do not sustain until the future point where follow-up data collection is conducted, rendering predictions unreliable.

Studies exploring prediction based on medical records assume regular and frequent contact with health professionals, which is not always the case for older adults, who may not seek help for symptoms they perceive to be related to the aging process (Smyer & Qualls, 1999). Furthermore, the fact that studies here report models with high predictive ability is not necessarily a reliable indicator that the models would be effective prospectively, since the majority of studies reported were retrospective in design. There is thus a need for exploration of approaches to data collection which involve monitoring over consecutive days, as well as more prospective studies to test models developed retrospectively.

The results also showed a large disparity between the number of studies predicting depression and the number predicting anxiety. It is possible that a greater number of attempts have been made to predict anxiety than were found in the searches, and that a positive publishing bias has meant that failed attempts to use machine learning to predict anxiety have not been published. The results may also suggest that there is greater difficulty in predicting future anxiety than there is in predicting future depression. Alternatively, the disparity may reflect a lack of interest in predicting future anxiety. It is important that research is conducted to explore these suggestions, and that results are disseminated to inform future approaches.

### Limitations

This study was conducted by one solo reviewer, and as such, results may be open to bias at the point of selecting studies to include in the analysis. Another limitation of the research is that masters and doctoral theses were not included in the searches. This may have meant that some studies that fitted the inclusion criteria were missed. A future review could include such types of publication for a more complete analysis of existing evidence, however the implementation of a quality assessment of papers would be recommended for studies of this type, to ensure that the studies were of equivalent quality to those presented in academic journals.

### Conclusion

A wide variety of machine learning tools have been applied to the problem of predicting future depression and anxiety, although there is a lack of research in applying such tools to older adult populations. Anxiety is less explored than depression, and research is needed to explore whether this is because anxiety is harder to predict than depression. An over-representation of retrospective studies means more prospective studies are required to better understand how different approaches may work in practice.

# Chapter 3 (B) Systematic Mapping Review 2

## 3.5 Introduction

### Background

Developing a successful tool for older adults to use on a digital technology platform requires an understanding of how older adults would relate to such technology, including what might motivate and prevent successful use of such a tool. A second mapping review of the literature was undertaken to explore older adults’ views of using technology to support mental health. Similar to the previous review, a mapping approach was taken to this review, since the aim was again to characterise the existing literature and establish research gaps. Booth and colleagues suggest that mapping reviews are suitable for this purpose (2016).

### Aims

This review sought to characterise the available literature on older adults’ views of technology to support mental health and to establish where there are gaps in knowledge. Stages in the review were the same as those in Figure 3 (Page 51).

### Research question

The question to be explored in this review was as follows:

* What are older adults’ views of using digital technologies to support mental health?

## 3.6 Methods

### Decision of search terms to include

Search terms were generated in the same way as in the previous review (see Section 3.2). Experts were consulted to refine the search terms. The experts were a professor of health services research and a research associate with experience in systematic reviews.

### Final search terms

Table 6 presents the main keywords used in the search. Exact search terms used in each database and dates of searches can be found in Appendix 2.

Table 6. Search terms used in the second mapping review

|  |
| --- |
| Over 50? or older adult? or older people or older person? or senior? or silver surfer? or people aged over 50 or elderly or elder? Or geriatric |
| AND |
| mobile phone application\*1 or cellphone application\*1 or cellular phone application\*1 or ipad application\*1 or tablet computer application\*1 or ipad app\*1 or tablet computer app\*1 or mobile phone app\*1 or smartphone app\*1 or smartphone\*1 or cellphone app\*1 or pc software or computer software or software or website\*1 or online or app\*1 |
| AND |
| perception? or experience? or attitude? or opinion? |
| AND |
| mental health or affective disorders or anxiety disorders or anxiety or major depression or depression or anxiety |

### Databases searched

The following databases were searched in this review. These were chosen based on relevance to the research question.

* PsycINFO
* ASSIA
* Web of Science
* MEDLINE via Ovid

### Inclusion criteria

Studies were included under the following criteria:

* English language, to reduce the time and expense associated with translating foreign language papers
* Discusses attitudes of participants aged 50 and over
* Published in the last ten years, to limit articles relating to out-of-date technology
* Discusses participants’ attitudes to applications, software or websites with a focus on mental health and wellbeing.

While the WHO and the NHS use 65 as the age at which a person becomes an ‘older adult’, here the decision was taken to include articles that discussed the views of people aged 50 and over. This was for two reasons – firstly, the number of relevant articles with participants exclusively over the age of 65 was very small, so including articles discussing samples with a lower age range increased the number of articles in the review. This increased the range of views for discussion, which was thought to strengthen the review overall. Secondly, the tool under development in this thesis is likely to take time to develop and test before it can be put to use. The time taken to implement a technology developed in research may be as long as fifteen years due to a range of complicating factors (Cresswell & Sheikh, 2013), by which time, those currently over the age of 50 will be over the age of 65. While age is likely to affect views of technology, a greater influence on views is likely to be the generation into which a person is born. The life experiences of current 65 year olds may not, therefore, be very relevant to the life experiences of 65 year olds in 15 years’ time, particularly in the field of technology where change occurs rapidly. Thus reviewing the views of people currently aged 50 and over is likely to be more informative of the life experiences of those people who will be over 65 when the tool described in this thesis is implemented.

### Search technique

This review used free text searching (Booth et al., 2016) to search the bibliographic databases listed above. Thesaurus functions were applied where available in each database.

### Sifting

Articles were sifted by title, abstract, then full article in accordance with the inclusion criteria. As with the previous review, where titles and abstracts were ambiguous, papers were kept in for the following round of sifting.

### Data extraction

Details were extracted to build an overview of the literature in the area. Data extraction involved collating the following details:

* Citation
* App(s)/website(s)/software(s) discussed
* Age range of participants
* Technology discussed (iPad, computer, laptop, mobile phone, etc)
* Methods used
* Summary of attitudes presented
* Notes

## 3.7 Results

The combined search results for all databases gave a total of 459 articles. 90 duplicates were removed, leaving 369 articles to sift. Figure 7 presents a PRISMA diagram of the results.

Records identified through database searching:

(n = 459)

Records screened:

(n = 369)

Full-text articles screened for eligibility:

(n = 12)

Studies included in the information extraction:

(n = 3)

Records after duplicates removed:

(n = 369)

Records excluded:

(n = 357)

Full-text articles excluded:

(n = 9)



Figure 7. PRISMA 2009 diagram showing selection of articles in Part B of the mapping review

Articles were first sifted by title. In this first sift, articles were excluded for the following principal reasons:

1. Not related to digital technology (152)
2. Not apps/websites targetting MH/depression/anxiety (99)
3. Attitudes to the technology/apps not discussed (8)
4. Not focussed on people aged over 50 (60)
5. Not an academic journal article (5)

A total of 324 entries were excluded, and 45 were taken forward for sifting at abstract level. In this sift, a further 33 articles were excluded for the following principal reasons:

1. Not related to digital technology (7)
2. Not related to depression/anxiety/wellbeing/mental health as a whole (or not technology related to these) (10)
3. Attitudes to the technology/apps not discussed (10)
4. Not focussed on people aged over 50 (6)

This left 12 articles for full paper review. On review of full papers, 9 further articles were excluded on the following grounds:

1. Does not discuss apps/websites targetting MH/depression/anxiety (1)
2. Not attitudes of people exclusively aged over 50 (4)
3. Not an empirical study (1)
4. No attitudes presented (3)

Three papers were finally included in the data extraction stage of the review. Extractions can be found in Table 7.

Table 7. Extractions from the systematic mapping review (Part B)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Citation | Participant age range | Condition targeted | App/website/software discussed | Technology discussed | Summary of attitudes |
| (Crabb et al., 2012) | 65-95 (mean 80.3) | Includes discussion on depression, anxiety stress and other mental health conditions | Use of the internet to get information on: healthy lifestyles, specific health information, drug/alcohol problems, depression, stress, anxiety, mental health issues, health behaviour and basic checks, organise appointments, set goals and measure progress, screen for MH conditions, learn new ways to manage stress/improve their mood | Any method of accessing the internet | Majority of respondents had used internet to search for health information. Few had experience of using the internet to source information about mental health. Majority said they would be interested in using internet for health-related purposes. |
| (Sauve et al., 2016) | 20 people aged 55-64 and 36 people aged 65+ | Physical, social, psychological wellbeing | Online educational game, "Live Well, Live Healthy" | "Computer equipment" | The educational game improved the users' perceptions of online games as a means of strengthening social ties and increasing social connectedness. The game decreased the number of participants reporting feeling depressed in the past week. The game increased reports of positive mood in the participants. |
| (Pugh et al., 2014) | 65 | Depressive symptoms | Online, therapist-assisted ICBT (internet Cognitive Behavioural Therapy) | Computer | The participant expressed contentedness with the system. He felt able to use the system, and felt it benefited him. |

### Characterisation of studies included

Very little literature is available on older adults’ attitudes towards technology to support mental health. This review found only three studies that had collected data from older adults on this subject, and in each of these cases, the collection of qualitative data on participants’ views of the technology was a subsidiary aim. The first of these studies (Crabb et al., 2012) involved a survey of 50 older adults aged 65 and over (mean age 80.3) which sought to explore, among other things, older adults’ internet use to obtain mental health-related information, and their interest in using internet interventions for health needs. Results showed that only a small number of participants (4 out of 50, 11%) had experience of using the internet to source information on mental health issues. Conversely, the majority of participants (39 out of 50, 78%) had an *interest* in obtaining information about health or mental health issues online. In addition, more than half of participants (32 out of 50, 64%) indicated an interest in using the internet to keep track of health problems.

There is likely to be some bias in this study since males were overrepresented in the sample (94% of participants were male), the majority of participants had higher than average incomes, and recruitment was conducted in just one region - the Silicon Valley region of California. These factors may have resulted in an over-reporting of positive attitudes towards technology, and greater experience of using it, since technology usage among the elderly is known to vary by socioeconomic status (SES), with those of higher SES being more experienced and comfortable using technology (Elliot et al., 2014; Vorrink et al., 2016). Male gender is also known to be correlated with greater internet use (Pew Research Centre, 2014).

Sauve et al. examined older adults’ perceptions about playing educational online games to learn about wellbeing (2016). Participants were 67 retired adults aged 55 and over. The study involved participants completing a set of questionnaires both before and after trying out a new online, multiplayer, educational bingo game called ‘Live Well, Live Healthy!’ over a period of one week. Findings showed that the experience of playing the game changed participants’ views of online games. Changes in views included a significant increase in the number of participants who believed that online games promote social connectedness (52% before, 66% after). Playing the game for one week also had a significant effect on the number of participants who reported feeling depressed in the past week – the percentage dropped from 17% before playing the game to 12% after. The results demonstrated that older adults’ attitudes towards apps and websites to support mental health are likely to change if they are given the opportunity to try using them. Findings may be limited by the small number of participants and the limited number of games which participants played.

The third study was an *n*-of-1 case study on the use of internet cognitive behavioural therapy (ICBT) for depression (Pugh et al., 2014). The form of ICBT discussed in the article involved the client working through materials on his own, and occasionally having email contact with a therapist. The study involved one male participant aged 65 who had depression. The case study presents highlights from email exchanges between the participant and his therapist, where he reflected on the experience of ICBT. The participant expressed many positive views about the programme, and reported particularly enjoying having contact with a therapist by email. While this article provides some endorsement for the provision of cognitive behavioural therapy for older adults using internet-delivered means, the generalisability of the study was limited by the extremely small sample. Furthermore, the participant was an experienced user of technology who was motivated to engage with ICBT, potentially biasing the views he expressed.

## 3.8 Discussion

To summarise the above results, existing research on older adults’ attitudes towards technology to support mental health is severely limited in quantity and quality. Just three studies have been identified in this area. Results of these studies indicate that older adults have generally positive views of technology to support mental health after they have tried using it. However, all of the studies are limited by small sample sizes, and all had potential for bias. This highlights a need for more research in this area.

Each of the above studies had only a subsidiary aim of understanding older adults’ views. Research designed from the outset to better understand older adults’ views of technology to support mental health would be valuable however. The use of technology holds promise for improving the health care of older adults, for example monitoring mood could hold promise for predicting future mental health problems (Astell et al., 2014), and older adults may be attracted to try new ways in which they could manage their own mental health without burdening others (Peek et al., 2015).

All studies in this review reported that their participants had positive attitudes towards the technologies and approaches described. Given the mix of both positive and negative opinions reported on older adults’ views of technology more generally (Greenhalgh et al., 2013; Chen & Chan, 2013), there is likely to be some bias in the studies mentioned here. This includes sampling bias for well-educated and high-earning participants in (Crabb et al., 2012). There may also have been bias in the selection procedure used by Pugh et al for their case study (2014). A full exploration of both positive and critical views of technologies to support mental health is needed to better understand how new mental health technologies could be developed and shaped to benefit older adults in future.

This is not the only review to explore the use of technology to support older adults’ mental health. Preschl and colleagues (2011) reviewed articles on different applications of technology to manage older adults’ mental health, including dementia. Crucially, their review covered only the effectiveness of these approaches and did not consider the views of the users themselves. Similar to the present study, they found the area to be poorly researched, and that existing studies are low in quality.

### Limitations

As with the first mapping review in this thesis (Chapter 3A), the present review was conducted by one solo reviewer, limiting the reliability of the screening process. In addition, the review excluded conference papers, masters dissertations and doctoral theses, meaning some studies may have been missed. Future work could involve a more comprehensive review which included these types of documents.

### Conclusion

A small number of studies have explored older adults’ views of technology to support mental health as subsidiary aims. These studies are prone to bias due to small sample sizes and unrepresentative recruitment strategies. More work is required to gain a full understanding of older adults’ views of using technology to support mental health, and this could involve new qualitative studies to explore issues around motivation to engage with such technology.

# Chapter 4 – An exploratory machine learning approach to the prediction of future depression

## 4.1 Introduction

This chapter discusses the application of a machine learning approach to secondary data on older adults’ mental state in order to generate predictions of future depression. It begins by examining reasons for using a machine learning approach, before exploring the differences between machine learning and statistical inference approaches. Next, a brief account is given of how the secondary data used in the analysis were collected. Consideration is then given to various machine learning techniques and their suitability to the problem in question. The methods chosen for this approach are then described, results are presented and then discussed.

### 4.1.1 Rationale

Machine learning has previously been applied to the prediction of future mental health conditions. Chapter 3A showed that these approaches have tended to involve either: developing models based on routinely collected clinical data (Jin et al., 2015); developing models based on data from MRI scans (Shimizu et al., 2015) or developing models based on mobile phone usage (Likamwa et al., 2013; Burns et al., 2011; Jiménez-Serrano et al., 2015).

None of these approaches is practical for the widespread monitoring of mental health in older people, however. Models derived from routine clinical data have only been applied in populations where frequent contact with health professionals is common, while many older adults fail to report symptoms of conditions they perceive to be part of ‘normal’ ageing (Smyer & Qualls 1999). Meanwhile, models based on MRI scans require the patient to undergo an MRI scan, which is costly for the healthcare provider and uncomfortable for the patient. Monitoring mobile phone usage assumes patients are regular users of mobile phones, and monitors users in a covert way, which may not be appropriate for older adults’ more sporadic use of mobile devices, particularly when older adults are known to have concerns about privacy in the use of digital technologies (Greenhalgh et al., 2013; Mitzner et al., 2010).

Easy to use self-report software that can be used on any platform may represent a more practical approach to detection of mental health problems within this population. Combining self-report data with a machine learning approach may offer the ability to successfully predict later incidence of a mental health problem, allowing early intervention to prevent crises. It is pertinent now to consider how a machine learning approach differs from a statistical inference approach to analysing data.

### 4.1.2 Differentiating machine learning and statistical inference

While both traditional statistical methods (hereafter termed ‘statistical inference’) and machine learning involve the application of mathematical tools to analyse datasets, there are some key differences between the two, although these are hardly discussed in academic literature. Here, three examples of differences are discussed: difference in aim, difference in methods and difference in culture.

#### Difference in aim

While the purposes of machine learning and statistical inference do overlap in some respects, there are some crucial differences between what they each try to do. Simply put, statistical inference aims to find out what is true in the world, while machine learning aims instead to make accurate predictions about the world.

Statistical inference approaches use a sample as a representation of a population. Statistical tests are then conducted on this sample to find out if some truth can be generated about the population. For example, a researcher may hypothesise that all right-handed people are also right-footed. By taking a sample of right-handed people and testing which foot is most dominant for each person within the sample, the researcher aims to collect enough evidence to validate the hypothesis about the footedness of all right-handed people – the whole population. The aim is to collect evidence in order to state a truth about the world.

Conversely, the aim of applying a machine learning approach is to develop an algorithm (a tool) for making predictions, without the need to state a prior hypothesis. For example, an estate agent may want to predict how much a house will sell for. Without stating any kind of hypothesis about which variables are likely to be predictive, the estate agent can use recent house sale data including for example the number of bedrooms, size of the garden and proximity to a train station, to develop a model to do just this. The aim is simply to produce a tool which can make accurate predictions, based on patterns found in past data.

#### Difference in methods

Many techniques can be used within both a statistical inference approach and a machine learning approach. These include, for instance, multiple regression and logistic regression. There are differences in how these same techniques are applied in the different approaches, however. These differences centre on the way that the validity of models is tested. In a statistical inference approach, assumption tests are used to pre-analyse the data on which models are built. In machine learning, fewer assumption tests are used while predictive accuracy tests are implemented after a model has been developed in order to examine the effectiveness of the model, for example using cross-validation. Cross-validation is a method of splitting a dataset and using one part of the data to train a model which is subsequently tested on another part of the data. Machine learning approaches may also involve deriving a model from one dataset, then using another dataset to test, or validate, the model. The train-test cycle in model development distinguishes machine learning by enabling good performance under conditions where statistical inference approaches may fail.

#### Difference in culture

Given that certain mathematical techniques are used in both statistical inference and machine learning, the idea of ‘culture’ becomes important – a different culture is associated with each approach (Breiman, 2001). This difference in culture is reflected in different terminology, for example a ‘sample dataset’ in statistical inference is referred to as a ‘training set’ in machine learning.

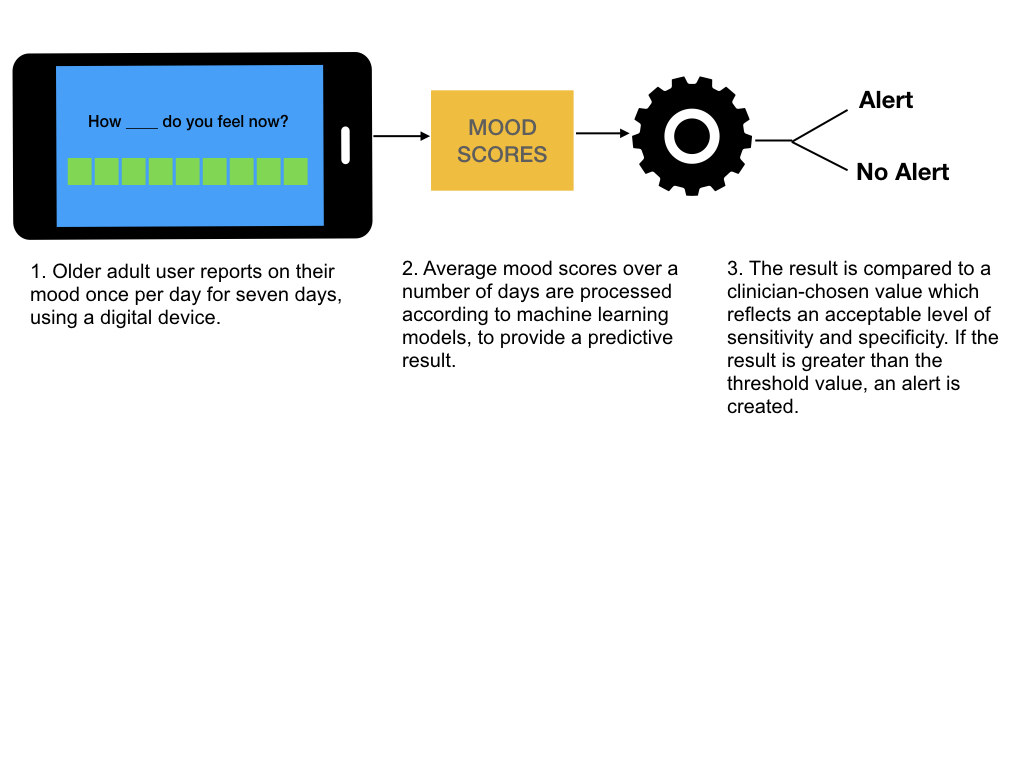
In mental health research, the tendency has been for modelling to be undertaken using statistical inference approaches, and this reflects the ‘culture’ in which mental health research has largely been conducted. However, this is beginning to change. The mapping review in Chapter 3A presents details of multiple studies which have applied machine learning techniques to the prediction of future depression and anxiety, indicating a widening interest in the application of machine learning in mental health research.

### 4.1.3 Summary

In summary, machine learning and statistical inference represent different cultures which typically seek to achieve different aims, and employ different methodologies. There is great potential for the application of machine learning techniques for the detection of mental health problems, and there are examples of research which have already applied machine learning techniques to this end.

### 4.1.4 Aims

The aim of this chapter was to explore the potential for developing a model for the prediction of future depression from older adults’ self-reported mood and appetite data. Chapter 3A showed that most input data in existing studies to predict future depression and anxiety were one-off measures, which could have been influenced by daily fluctuations. Thus, the present work sought to use more dynamic measures to record data over a longer time period. Figure 8 indicates how collecting and processing data in this way could be used to provide alerts to promote intervention before crises occur.



**Figure 8. Diagram showing potential application of a machine learning approach**

It is now pertinent to introduce the secondary dataset used to explore the development of machine learning models of mood and depression.

## 4.2 The NANA dataset

The Novel Assessment of Nutrition and Ageing (NANA) homesystem (Astell et al., 2014) was developed to capture reliable and valid data from older adults in their own homes without a researcher or clinician being present. NANA is a multidimensional toolkit that can be used to assess the diet, mood, appetite, cognition and physical activity of older adults on a daily basis. NANA includes measures of mood and appetite developed by Brown and colleagues (Brown et al., 2016). The toolkit permits digital collection of data using a touchscreen interface. Figure 9 shows the NANA homesystem software being used to collect mood data on an Asus Eeetop computer with 15 inch screen.



**Figure 9. Older adult self-reporting mood using the NANA Homesystem software on an Asus Eeetop 15 inch touchscreen PC**

### 4.2.1 NANA data collection procedure

A validation study of the NANA system was conducted in 2011 (Astell et al., 2014), and the study provided a rich dataset which is used as the basis for this study. Forty older adults were provided with a NANA homesystem to use, 20 in Sheffield and 20 in St. Andrews. The trial was conducted over an 11-week period, although participants only used the NANA system for three of these 11 weeks, with a three week ‘washout’ period between each week of data collection. Pen and paper tests of cognition and depression were taken in the first week of the study, and these were repeated in the last week. Adults in the study were aged 65-89 (Astell et al. 2014). Figure 10 shows the study design for the NANA validation trial, indicating the time periods where participants used the NANA homesystem.

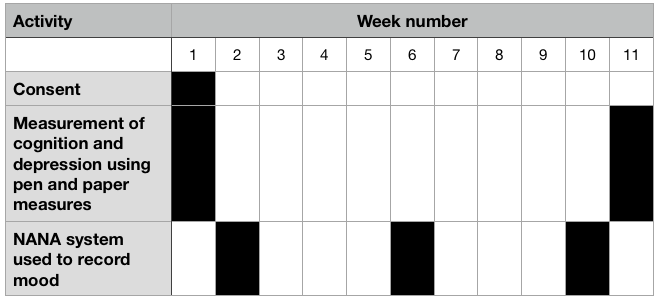


Figure 10. Study design for the NANA validation study. Adapted from (Astell et al., 2014)

The NANA system was comprised of a touchscreen computer with a 15-inch screen and a camera, which participants used to record details of everything they ate and drank during data collection periods. In addition, on a daily basis, participants were prompted to use the NANA system to take two tests of cognition, repeated three times consecutively each day. Users of the NANA system were also prompted to report on their mood and appetite once daily by answering six questions that appeared on the screen. These questions were of the format “How xxxxxx do you feel right now?”, where xxxxx was replaced by each of the six adjectives listed in Table 8 alternately in a random sequence. Users answered each question by pressing a button to score their experience of each feeling on a 0 to 10 scale where the screen informs them that 0 is ‘Not at all’ and 10 is ‘Very’.

Table 8. Mood and appetite descriptors shown to participants on the touchscreen in the NANA validation trial within the question "How \_\_\_\_ do you feel right now?". Participants responded by selecting a number between 0 and 10.

|  |
| --- |
| Happy |
| Sad |
| Relaxed |
| Tired |
| Hungry |
| Alert |

Gold standard pen and paper measures of depression and cognition were taken at the beginning and end of the 11-week data collection period. The depression measure used was the GDS (Yesavage & Sheikh, 1983). The GDS has two versions, a 30-item version and a 15-item version (Yesavage & Sheikh, 1983; Sheikh & Yesavage, 1986). In the NANA validation study, the 15-item version was used. In this version, a score of five points or more is indicative of depression (Sheikh & Yesavage, 1986). However, since the publication of the scale, some have argued for using a cut-off of three points (Arthur et al., 1999). In either case a high score on the GDS does not on its own constitute grounds for a clinical diagnosis of depression, rather the scale gives an indication of whether further assessment is required.

### 4.2.2 Appropriateness of the data for predicting depression

Depression is a condition characterised by tiredness, loss of pleasure in life, depressed mood, unclear thoughts and changes in appetite (see Chapter 2). As such, the mood and appetite measures in NANA (ratings of alertness, happiness, hunger, relaxation, sadness and tiredness) may be sensitive to depression and/or its prodrome. Here, machine learning is used to explore the potential for using the mood and appetite measures to predict future occurrence of depression.

## 4.3 Descriptive exploration of the data

A descriptive exploration of the data was undertaken to achieve a good basic understanding of the data and to inform the choice of techniques to apply. While the data resulting from the NANA validation study included many different measures, here only a subset is explored. First, the demographic information about the participants is explored. Second, scores on the GDS for all participants are presented, along with the distribution of mood and appetite scores from throughout the validation study.

### 4.3.1 Participant demographics

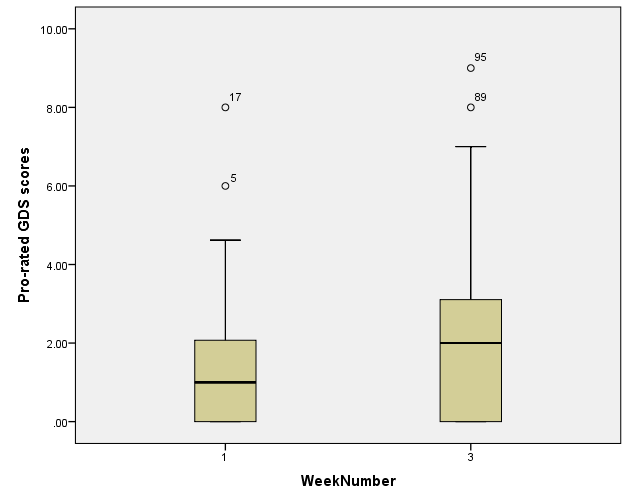
The NANA system was installed in the homes of 40 older adult participants, of whom 16 were male. The participants ranged in age from 65 to 89, with a mean age of 71.88.

### 4.3.2 Scores on the GDS

Figure 11 shows the distribution of scores on the GDS at baseline and follow-up. Across the group, scores on the GDS increased slightly during the study from a mean of 1.62 to a mean of 2.21 between the start and the end of the trial (see Table 9), although this increase was not significant [for details, see (Astell et al., 2014, p.104)]. Five participants who had scored under the cut-off of five points at baseline scored above it at nine weeks follow-up, while four participants had an increase from below a cut-off of three at baseline, to above a cut-off of three at follow-up (see Table 10).

Table 9. GDS summary statistics from the NANA validation trial

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Timepoint | Mean | Median | SD | Minimum | Maximum |
| Baseline | 1.62 | 1 | 1.89 | 0 | 8 |
| Follow-up | 2.21 | 2 | 2.43 | 0 | 9 |
| *Note.* Scores out of 15, pro-rated.  Table 10. Frequencies of participants in the NANA validation trial valued as depressed according to the GDS at cut-offs of three and five points (15-item)   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **GDS cut-off used** | No. scoring over the cut-off at baseline | No. scoring over the cut-off at follow-up | No. scoring below cut-off at baseline, but above cut-off at follow-up | No. scoring below cut-off throughout the study | | **≥3** | 9 | 13 | 4 | 23 | | **≥5** | 3 | 7 | 5 | 32 | | | | | | |



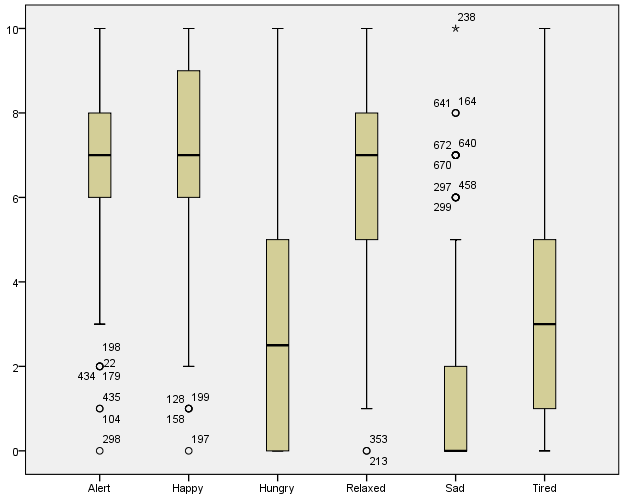
1 2

**Time point (T)**

Figure 11. Boxplot showing scores on the GDS at baseline (T1) and at follow-up (T2) in the NANA validation trial (Astell et al. 2014)

### 4.3.3 Mood scores

The overall pattern of mood and appetite scores showed higher scores to be commonly reported for positive affect items (‘Alert’, ‘Happy’, ‘Relaxed’) and lower scores for negative ones (‘Hungry’, ‘Sad, ‘Tired’), as shown in Figure 12. Scores for ‘Hungry’ had the greatest interquartile range, indicating greatest variability in the data for this item, while scores for ‘Sad’ and ‘Alert’ showed the least (Figure 12).



**Score**

**Mood word**

Figure 12. Variation in participants' scores given for different mood words across all days in week 1 mood reports. Participants rated their mood on six words (Alert, Happy, Hungry, Relaxed, Sad, Tired) once daily.

## 4.4 Methods

Several different approaches were used to study whether the data collected on an everyday basis could be used to separate those who would score above a threshold for depression on the GDS at follow-up, from those who would not. These approaches included both supervised and unsupervised learning methods to derive algorithms for automatic classification. Supervised learning methods use an outcome measure to group data points, then analyse the data points in each group to determine within-group similarities. These similarities can then be used as rules to classify new data into appropriate groups, providing predictions of the outcome measure. By contrast, unsupervised learning methods analyse all data points in a dataset and find similarities by which to group them. These groupings can then be used to classify new data.

Though an outcome variable was available in this dataset (GDS scores), the decision was taken to explore unsupervised techniques as well as supervised techniques, to identify patterns in the structure of the mood data collected. The choice of which techniques to apply was informed by literature on uses of different machine learning techniques. This section first explains how the data were prepared for analysis. After this, an account is given of the supervised and unsupervised learning techniques used with this data. The benefits and shortcomings of each technique are explained and results are provided for the final model development.

### 4.4.1 Data cleaning

Data cleaning involved removing lines of data in one of three cases: (i) the participant did not complete the follow-up GDS test (ii) the participant gave answers to one of the mood questions on fewer than three days (iii) the participant completed two or more days’ mood questions within 15 minutes of one another. The reasons for these removals were as follows: (i) if the participant did not complete the second GDS test, it was not possible to use supervised learning methods as the second GDS test was the outcome variable. (ii) Mood scores were averaged across the week. If participants had only reported mood on one or two days, it was more likely that anomalous results could skew the data. (iii) Participants were prompted to report on their mood once per day. If a participant did not complete the mood report within the day, the prompt would remain on the system, and would ‘stockpile’ with the following day’s prompt. As such, when participants caught up and completed these reports, it was not known if they were trying to recall how they had felt the day before (unlikely to be accurate) or if they were truly reporting how they felt on the day. As such, these records were deleted.

Thirty-seven entries remained when the unusable data had been removed listwise (i.e. all data for a participant was removed if a crucial part was missing, as described above). Listwise deletion was necessary in this case because each technique used in the machine learning approach depended on the use of a full set of variables for each participant.

### 4.4.2 Cluster analysis

Cluster analysis is an unsupervised learning tool which separates data in a multidimensional space according to the properties of each data point. It then collates data points which are similar to each other. Figure 13 shows this idea in a visual format. In this example, two-dimensional data are distributed in two clusters, indicated by the green and red circles. Cluster analysis numerically separates these clusters, and assigns a group identity to each data point. This method also works in datasets with more than two dimensions. Since this process does not use a dependent variable to determine groupings, this is considered an unsupervised learning technique.

Figure 13. Scatter graph showing an example of clustering. The green and red circles illustrate how cluster analysis separates data points into different groups.

The cluster analysis technique was applied to the mood and appetite data from the NANA dataset to determine if it would separate individuals in a way which was indicative of whether each participant would go on to score above the cut-off for depression on the GDS. The tool separated the data into two clusters, which were highly sensitive to the later onset of depression - all the individuals who later scored above the cut-off were assigned to the same group. However, the cluster solution provided a low level of specificity – many participants who did not go on to score above the cut-off were also grouped with those who did. Due to this low specificity, this technique was not likely to be of use to clinical decision-makers, since many people would be incorrectly assessed as likely to become depressed. Therefore, full results are not presented here.

After applying the clustering technique described above, and finding that it produced a model with poor specificity, the decision was taken to apply a supervised learning technique to the NANA study data. It was important to use a technique that would be applicable to a small number of individuals, since there were only 37 individuals within the cleaned dataset, with only seven experiencing depression at follow-up. Many machine learning techniques require larger datasets with a greater proportion of positive cases to work effectively. The small number of individuals in the NANA study limited the number of approaches that were possible in this exploration. The first supervised technique applied to the data, which was determined to function well with a smaller number of individuals, was regression.

### 4.4.3 Regression

The term regression describes a process used to approximate a trend line in a multidimensional space, given a set of multidimensional data points. This line can also be called a regression line. Regression lines are often superimposed on scatter graphs. See Figure 14 for an example.

Figure 14. Scatter graph with trend line. Demonstrates linear regression within a two-dimensional space.

In this example, a straight line is produced, and this type of regression is called linear regression. To generate this regression line, the linear equation for a straight line which passes as close as possible to each of the plotted data points is calculated. Least squares regression is one method for achieving this. This was first used in 1795 by Gauss (Aldrich 1998). The method applies a mathematical formula to find the line where the sum of the squared distances from the line to each data point is minimized. Once the formula for the line has been derived, this formula can be used to predict a value of y (an outcome variable) for any given value of x (input variable).

It is relatively easy to visualise the application of linear regression in two or even three dimensions. Though less easy to visualise, it is possible to apply this technique to problems with four or more dimensions, where a new dimension is required for each independent variable. This is termed ‘multiple regression’. In the mood and depression scale data from the NANA study, there are seven variables – six input variables (mean averaged scores on each participant’s daily happiness, sadness, alertness, tiredness, relaxation and hunger) and one outcome variable (follow-up GDS scores). Applying multiple regression to these data could generate models allowing prediction of follow-up GDS scores, given new data on mood from similar older adults.

However, rather than predicting scores on a scale, it is likely to be more useful for clinicians to know whether a patient is likely to score above the cut-off for depression or not, as this provides information about whether a patient requires intervention of some kind. This kind of information can be summarised in binary form, where a value of one is equal to a positive case (requiring intervention) and a value of zero is equal to a negative case (not requiring intervention). Logistic regression can be used to generate models for the prediction of variables of this type.

### 4.4.4 Logistic regression

Logistic regression is similar to linear regression in that it approximates a line of best fit. However, logistic regression uses a logit transformation, creating an s-shaped curve rather than a straight line. Points on this curve represent the probability that a binary outcome will be positive (1 rather than 0). Like multiple regression, this can be used in a multi-dimensional space.

Logistic regression is a supervised learning technique which uses a binary outcome variable. In the NANA data, the outcome variable used is the GDS, which is a scale scored out of 15. This represents continuous rather than binary data. However, the instructions for the 15-item GDS scale indicate that a score of five or more suggests depression (Yesavage & Sheikh, 1983), and more recent work has also advocated using a cut-off of three points (Arthur et al., 1999). As such, it is possible to transform scores on the GDS into binary data by coding scores above the cut-off as ‘1’ (depressed) and scores below the cut-off as ‘0’ (not depressed). Logistic regression can then be used to train a model for the data, using self-reported mood scores (participant mean scores for happy, sad, tired, relaxed, alert and hungry) as the input variables and GDS depression status as the outcome variable. The model could then be used to estimate whether patients newly reporting their mood and appetite using the six self-report questions in NANA would score above the cut-off for depression at a later date.

A common problem with regression modelling is that of overfitting. Overfitting occurs when a model fits the sample data too closely, without regard to its inter-sample behaviour. Some techniques in machine learning can be used to avoid overfitting. Here, two techniques were explored and one was chosen to use in the final analysis.

### 4.5.1 Ridge regression

Ridge regression is a method which controls variance by regularizing the coefficients in a regression model. This has the effect of reducing the prediction error of the model when applied outside the sample. When ridge regression was applied to the NANA data, some of the coefficients were reduced to very low levels, and this meant that including some of the variables did not enhance the model, while still adding to the burden of data collection (i.e. the user of the software would have to report on all six of the mood variables when only a minority of these made sizeable contributions to the model). Therefore, a different technique was sought which would improve parsimony by excluding variables that contributed least.

*N.B. The remainder of this chapter is in large part taken from a publication authored by the researcher and his supervisory team (Andrews et al., 2017).*

### 4.5.2 Selection of variables using the LASSO

Similar to ridge regression, the Least Absolute Shrinkage and Selection Operator (LASSO) (Tibshirani, 1996) is an operation that introduces bias into the estimated coefficients of a (here, logistic) regression procedure by shrinking their absolute size (Tibshirani, 1996). This procedure prevents overfitting and provides knowledge of which variables are most informative. In contrast to ridge regression, LASSO omits coefficients for those variables which contribute little to the model, resulting in a more parsimonious solution.

To find the best performing model, a “shrinkage” parameter, lambda, must be set, and this was chosen using a repeated stratified cross-validation framework – the currently accepted best practice in machine learning. Cross-validation permits the use of all data in estimation and testing of models and guards against the tendency of more complex models to overspecialise. This is especially important when there are a relatively large number of variables and a small sample size as is the case here. In *k*-fold cross-validation, the data are divided into *k* non-overlapping sets. At each stage, one set is held-out for comparison, while a model is fitted to the remainder. This is repeated for each of the *k* holdout sets. In this study, cross-validation was used for parameter selection, setting *k*=5 (larger values would not allow each hold-out fold to contain a data item with a positive outcome). In addition, two hundred Monte Carlo repetitions were made of the cross-validation procedure, to guard against the possibility of a favourable random 5-fold partition.

One hundred different values of lambda were used within the procedure to determine the level at which the usual goodness-of-fit measure, deviance, was minimized. Then the greatest value of lambda that lay within one standard error of this point was chosen to use in the model. This technique has been used by Breiman among others, and is justified on the basis of parsimony (Breiman et al., 1993; Hastie et al., 2009; Krstajic et al., 2014). The increase in deviance associated with this method is negligible. Choosing the largest possible value of lambda has the effect of selecting out variables which make no contribution, or only a minimum contribution, to the model.

### 4.5.3 ROC curve analysis

When producing a new diagnostic or predictive tool, an ROC curve analysis can be used to assess the predictive ability of the tool. An ROC curve analysis was used in the present study to assess predictive ability of the two models generated with LASSO (one using a cut-off of three points on the GDS, and one using a cut-off of five). This section describes what an ROC curve is and how it is used.

Diagnostic and predictive tools have varying levels of sensitivity and specificity. Sensitivity describes how likely it is that a positive case will be correctly identified. For example, if two in 30 people contract malaria, how likely is it that a new blood test will detect the disease in both individuals? This can also be described as the ‘true positive rate’. Specificity describes how likely it is that only true positive cases will be deemed to be cases. If two in 30 people contract malaria, how likely is it that a blood test will flag only these two individuals?

An ROC curve is a two-dimensional graph which plots sensitivity against (a transformation of) specificity, to visually represent the performance trade-off between the two at different cut-off values of the predictor. While the line x=y on an ROC graph indicates a chance level of prediction, an ROC curve for a classifier typically bends towards the top left corner of the graph. The further it bends in this direction (the closer it is to the top left corner) the higher the possible combination of sensitivity and specificity that can be achieved simultaneously.

ROC curves are used in practice for clinicians to decide where to set a threshold for an action to be taken, given the potential trade-off between sensitivity and specificity at each threshold value. For example, a clinician might consult an ROC curve and decide that a threshold where sensitivity is high (e.g. 0.8) but specificity is relatively low (e.g. 0.5) may provide a desirable test for early signs of cancer, because the cost of seeing more patients who may not have the disease is outweighed by the benefits of catching a small number of true cases early on.

Literature on research using ROC curves reports the area under the curve (AUC) as a figure between 0 and 1. This value is equivalent to the chance that the predictive model will rank a randomly selected positive case higher than a randomly selected negative case. The predictive ability of different models can be compared using the AUC.

## 4.6 Results

The first model derived using the LASSO used a cut-off of five points on the GDS. Figure 15 shows the deviance plot for this model, on which the green vertical line indicates the point where the goodness-of-fit measure, deviance, was minimized. At this point, the value of lambda was 0.062. The blue dotted vertical line in Figure 15 indicates the point with the largest value of lambda where the deviance fell within one standard error of the minimum. Here, the value of lambda was 0.075. This value of lambda was used for the subsequent steps in the process.

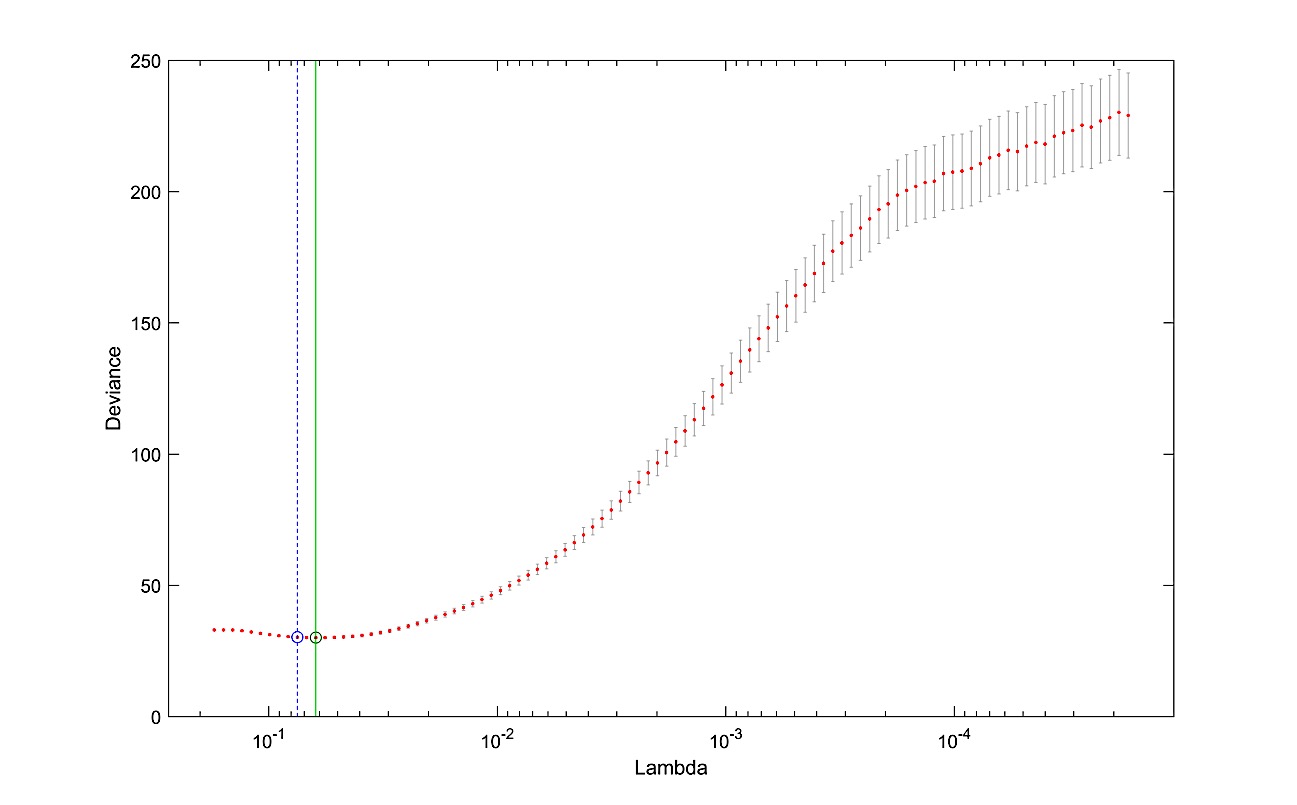


Figure 15. Deviance plot showing deviance at different values of lambda for the model with a cut-off of five points on the GDS.

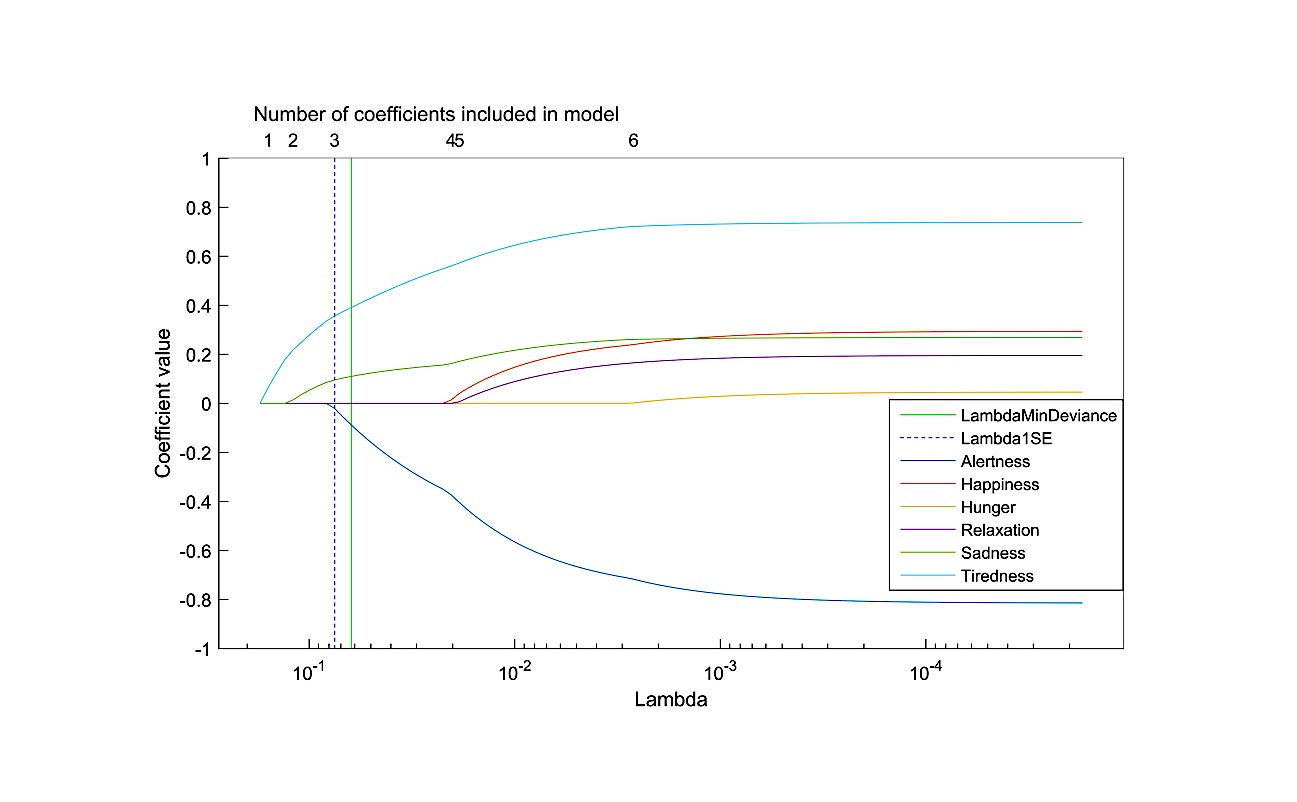


Figure 16. Coefficient plot showing contribution of each coefficient to the logistic model.

Figure 16 shows a coefficient plot for the model using a cut-off of five points on the GDS. The coefficient plot compares the contribution of each of the input variables (alertness, happiness, hunger, relaxation, sadness and tiredness) to the model, at different values of the shrinkage parameter, lambda. The green line indicates the value of lambda where the deviance was minimized, while the blue dotted line indicates the largest value of lambda which fell within one standard error of the minimum. At this value of lambda (0.075), three variables had non-zero normalised coefficient values in the model (sadness, tiredness and alertness).

The variable contributing least to the model (‘alertness’) has a relatively small normalised coefficient value (-0.0211). Thus, this variable was removed from the model, as the practical advantage in reducing the demands on the client group was judged to outweigh the cost of a further increase of deviance. These choices resulted in less than a 1% overall increase in deviance. The final model retains only a constant (-2.99) and the coefficients for ‘sadness’ (0.09) and ‘tiredness’ (0.34), corresponding to the largest value of lambda that excludes ‘alertness’ (0.082). Coefficients for sadness and tiredness were both positive, suggesting that higher average sadness and tiredness scores over Period 1 were associated with higher risk of scoring positively for depression at follow-up GDS in this sample.

The results of the ROC analysis for this model can be found in Figure 17. The AUC for the LASSO with logistic regression was 0.88.

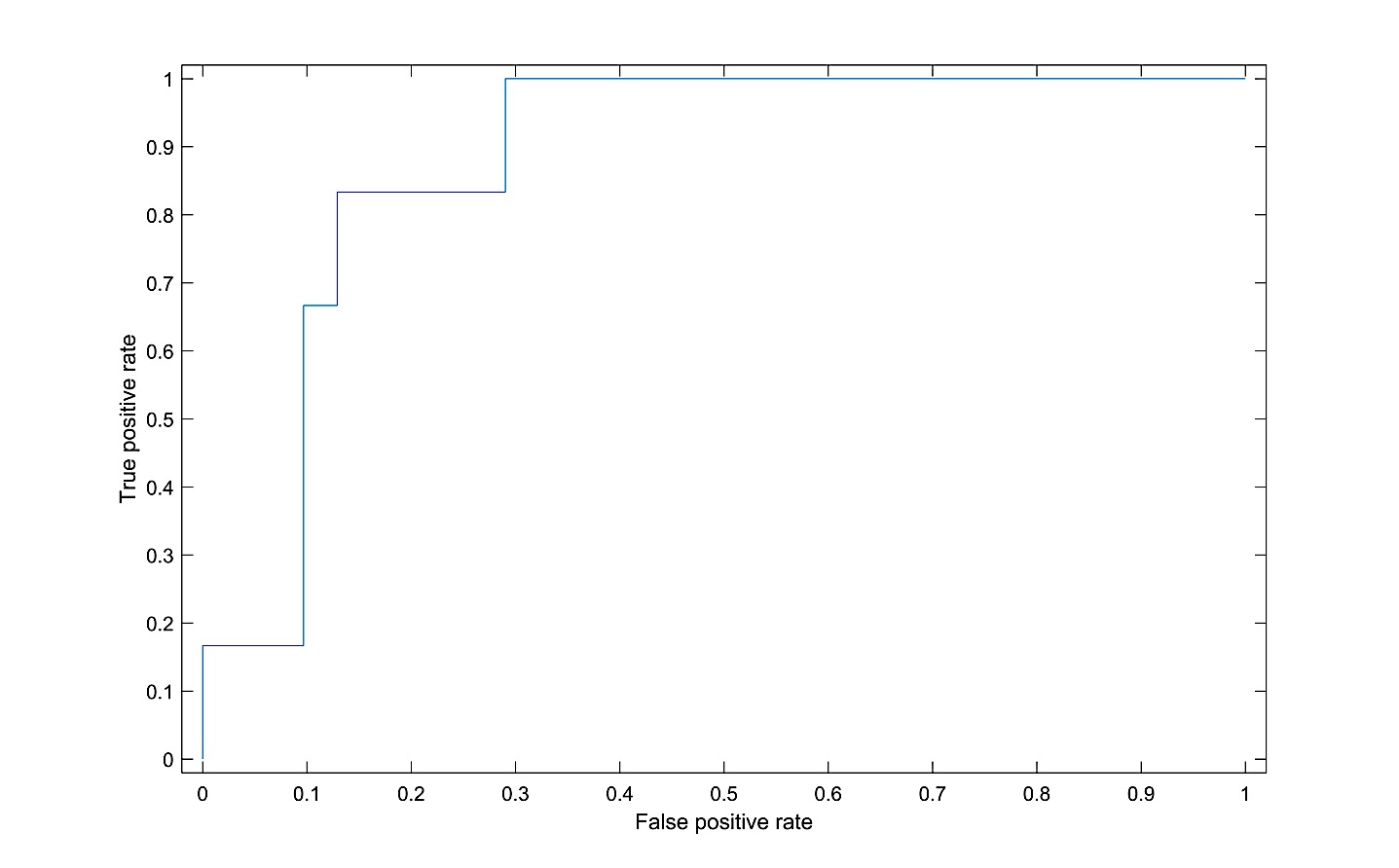


Figure 17. Receiver Operating Characteristic curve for predictive model derived using LASSO for the model with a cut-off of five points on the GDS. The AUC is 0.88.

Next, a model was derived using a cut-off of three points and above on the GDS as a positive case of depression. Here, the deviance was minimised with a lambda value of 0.044 (Figure 18). The largest value of lambda within one standard error of the point where deviance was minimized was 0.058.

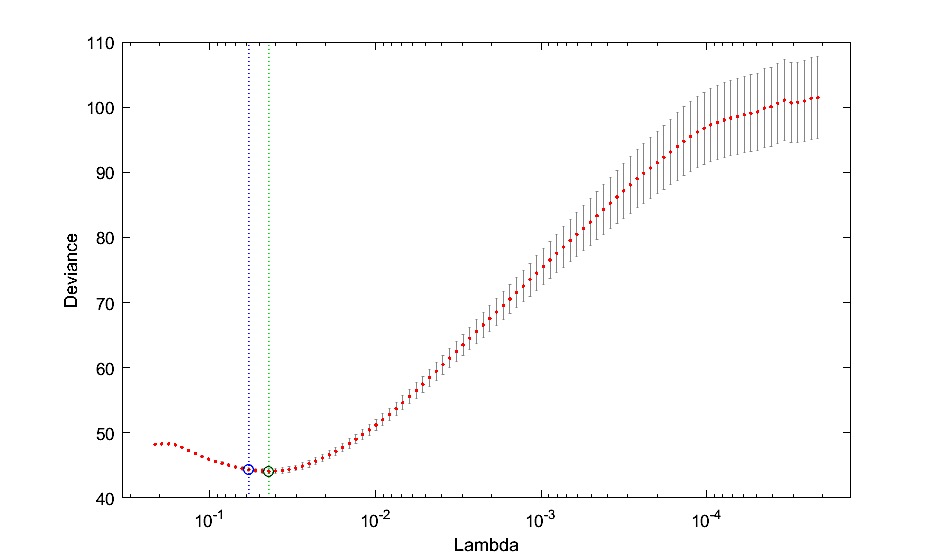
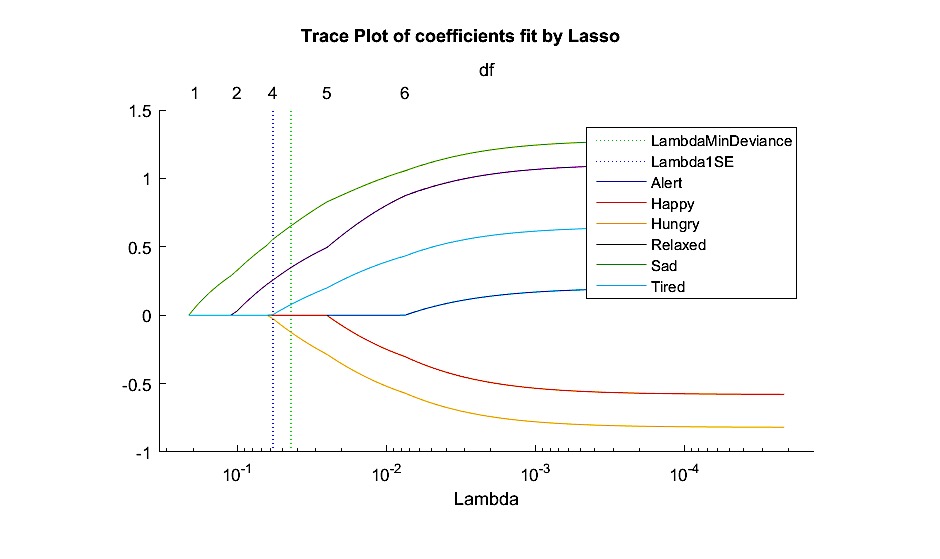
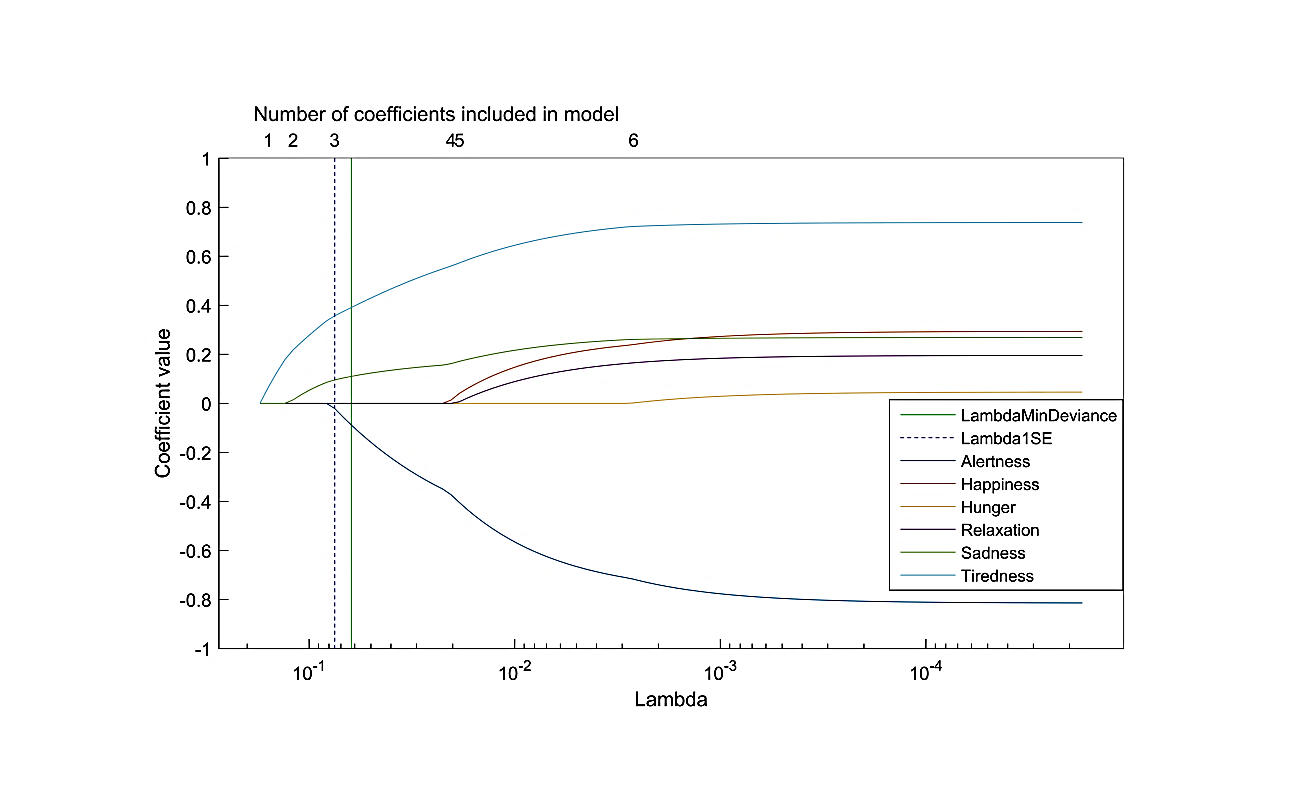


Figure 18. Graph showing deviance plot for model using a cut-off of three points on the GDS.



Number of coefficients in the model

Coefficient value

Figure 19. Coefficient plot for model derived from using LASSO with the NANA validation data with a cut-off of three points on the GDS

Figure 19 shows the coefficient plot for the model derived using a cut-off of three points on the GDS. At the value of lambda chosen above (0.058), the model included four variables, (‘sad’, ‘relaxed’, ‘tired’ and ‘hungry’) as well as a constant (see Figure 19). Normalised coefficient values for ‘hungry’ and ‘tired’ were small relative to the other values, so these were omitted, causing a small increase in deviance. Figure 20 shows the ROC curve for the resulting model, where the AUC is 0.80.

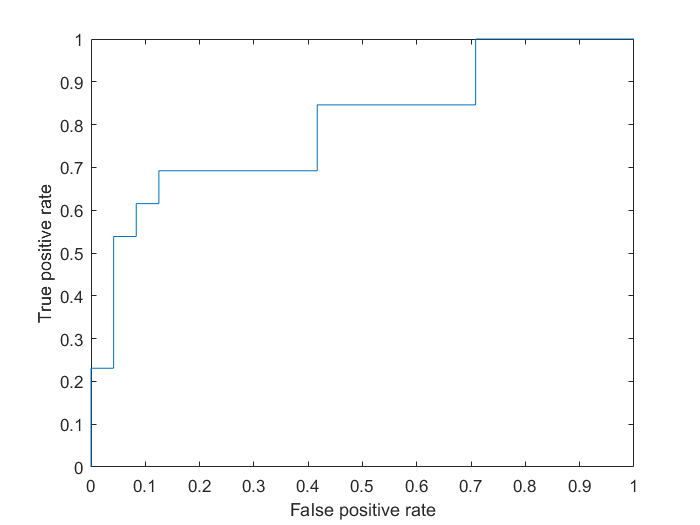


Figure 20. Graph showing ROC curve analysis for model derived from the NANA validation data with a cut-off of three points on the GDS. The AUC is 0.80.

## 4.8 Discussion

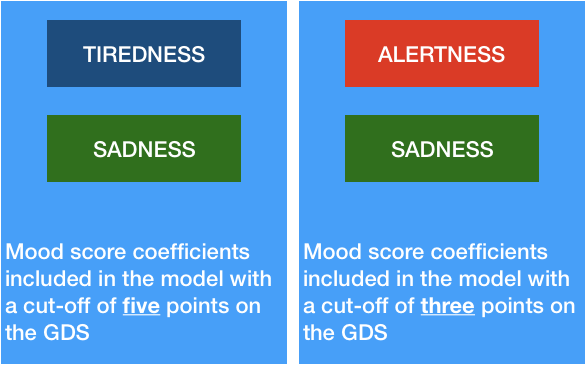
The LASSO method with logistic regression was used to produce a predictive indicator of depression status based on data from touchscreen mood self-reports of older adults. Two cut-offs on the GDS were explored. The model with a cut-off of five produced an AUC of 0.88, indicating good predictive ability within this sample. The model with a cut-off of three produced an AUC of 0.80. While the AUC is a good portmanteau measure of performance, a particular clinical setting would require a choice of operating threshold at which to make a decision. The ROC curve can be used to determine, for example, the acceptability or otherwise of the sensitivity (true positive rate) associated with a chosen ‘false alarm rate’ (false positive rate or one-minus-specificity).

Where previous research has included ROC analyses as a measure of predictive ability, it is possible to compare the performance of their models against those presented here. Table 11 shows references and AUCs for those papers in the systematic mapping review in Chapter 3A which reported a value for the AUC. The AUCs for the two models presented here were 0.88 for the model with a cut-off of five points on the GDS and 0.80 for the model with a cut-off of three. The papers referred to in Table 11 report AUCs of between 0.63 and 0.84. In these terms, the AUCs for the models presented here are comparatively high.

Table 11. Table showing AUCs for models produced in other approaches to predicting future depression

|  |  |  |
| --- | --- | --- |
| **Paper reference** | **Machine learning technique applied** | **AUC** |
| (Dabek & Caban, 2015) | Neural network model with hidden layer and SVM layer | 0.76 |
| (Jimenez-Serrano et al., 2015) | Naive Bayes, logistic regression, support vector machines, artificial neural networks | 0.66 – 0.74 |
| (Jin et al., 2015) | Logistic regression with ridge parameter, multi-layer perceptron, support vector machine, random forest | 0.78 – 0.81 |
| (Kessler et al., 2016) | Ensemble regression trees, LASSO, elastic net, ridge regression. | 0.63 – 0.76 |
| (Mourao-Miranda et al., 2012) | Gaussian process classifiers | 0.78 |
| (Tortajada et a al., 2009) | Multi-layer perceptron (artificial neural network) | 0.80 – 0.84 |

The LASSO procedure is designed to reduce the number of variables necessary to produce a model, and to optimize the size of the coefficients included in the model. Of the six mood variables used here (alertness, happiness, hunger, sadness, tiredness, relaxation), two were retained in each of the final models (sadness and tiredness at cut-off five, sadness and relaxation at a cut-off of three, see Figure 21).



**Figure 21.** **Diagram showing the coefficients included in the two models developed using the LASSO**

These results suggest that a higher average score for sadness was indicative of later positive depression status in both models. Tiredness also featured in one of the models, and this may reflect participants experiencing a lack of energy. The Diagnostic and Statistical Manual of Mental Disorders 5 (American Psychological Association, 2013) states that low mood and lack of energy are symptoms of depression. These findings would suggest that in this sample, these symptoms could be detected early in the course of a depressive episode using a daily self-report.

While some research has already explored the application of machine learning to the prediction of future mental health conditions, the work presented here represents a new contribution to the field. Although the LASSO has previously been used to predict future depression (Brann et al., 2017; Wardenaar et al., 2014), it has not been used within a population of older adults. Furthermore, the use of self-report data collected daily as the basis for machine learning work also represents a divergence from existing approaches to the prediction of future depression. These have previously relied on the use of medical records (Bian et al., 2017; Dabek & Caban, 2015), one-time self-report scales (e.g. Rude et al., 2010; Schalinski et al., 2016), or neurological measures (e.g. Long et al., 2014; Foland-Ross et al., 2015). Given the comparatively high AUCs reported here, the results are indicative that this novel approach could hold promise for future application in healthcare, if validated prospectively.

### 4.8.1 Strengths and limitations

This study has a number of strengths. Using a LASSO procedure optimizes the structure of the model by including only a limited number of variables. This creates parsimony and reduces the burden of data collection on users. Use of cross-validation creates a ‘closed loop’ in the development of the model, which focuses attention on predictive ability rather than inference. By applying a machine learning approach, a model has been generated with the potential to predict depression status in older adults.

There are limitations to the study. The final model produced in this study has been trained on just 37 individuals from two British cities. In its current form it is therefore unclear how well it would apply in clinical use, though the aim of this study was not to produce a clinical tool. Rather, it has demonstrated a technique which could be used with a larger and more representative sample to produce a model, which could in turn be put into clinical use with people at risk of a mood disorder. Accurate prediction of such a condition could allow early intervention in line with national or international guidelines, for example those offered by the World Health Organisation in their Mental Health Gap Action Programme intervention guide (WHO, 2010), potentially reducing the number of older adults who develop a mood disorder and/or commit suicide.

There are also limitations to the analysis used. The model is tested on the original sample so results are open to criticism. However, the use of repeated 5-fold cross-validation serves to ameliorate the problem of over-fitting and is known to offer conservative results (Rodriguez et al., 2010). This suggests that, in prospective use, performance might exceed that quoted.

## 4.9 Conclusion

Application of machine learning appears to be appropriate for self-report mood data. This is the first time that the logistic model, in conjunction with the LASSO, has been used to produce an algorithm for prediction of depression status based on older adults self-reporting their mood using a touchscreen interface. The comparatively high predictive ability of the models provides evidence to support the approach. Further work could validate the models produced here prospectively.

# Chapter 5 - Older adults’ perspectives on technology to report on mood and support mental health

## 5.1 Introduction

This chapter presents a qualitative exploration of older adults’ perspectives on the use of technology to report on mood and support mental health. The chapter begins by situating the study within the context of the thesis. It then sets out the research questions and approach to the study, before describing the methods used, including how these were developed. The findings are then presented, with explorations of participants’ views under each of the main themes derived in the analysis. These sections are followed by a discussion, where each of the research questions are addressed in light of the data, and the study is critically evaluated.

### 5.1.1 Background

This thesis aims to explore how machine learning might be applied in practice to the prediction of future depression and anxiety in older adults. The approach chosen requires participants to report on their mood over a period of seven days using a digital device. For such an approach to be applied in practice, it must address usability requirements specific to older adults. Furthermore, older adult users must be motivated to use it every day (over a time-limited period). It is important to explore older adults’ views on the use of such a tool, to understand the potential for its widespread use, as well as practical aspects of the use of such a tool.

Section 2.6 of this thesis explored literature on older adults’ attitudes towards technology in general. This included findings that older adults enjoy using technology to support their healthcare, for example being able to order prescription refills by phone, measure blood pressure at home, monitor their weight, and research health conditions online (Mitzner et al. 2010). Section 2.6 also described research exploring barriers to older adults’ use of technology. These included that older adults perceive technology to be expensive (Greenhalgh et al. 2013; Chen & Chan 2013), clunky or oversized (Greenhalgh et al. 2013; Peek et al. 2015), and that often, older adults prefer the old-fashioned ways of doing things (Peek et al. 2015). Prior research has also highlighted that older people can have trouble remembering how to use new technologies (Chen and Chan. 2013; Peek et al. 2015).

Chapter 3(B) systematically reviewed the literature on older adults’ attitudes to technology to support mental health, and found very few papers on this topic. Those that were identified presented overwhelmingly positive responses to the technology participants used, for example participants’ positive views of the benefits of online games for mental health increased by a statistically significant amount during one of the studies (Sauve et al. 2016), and in another, the only participant in the study reported being very satisfied with the online, therapist-assisted CBT that he received (Pugh et al. 2014). Methodological weaknesses may have meant this positivity reflected bias in these studies. Research in (Crabb et al. 2012) reported that very few older adults (4/50) had used the internet to source information about mental health issues, although more than half of the sample (56%, 28/50) expressed an interest in using the internet to source information on health or mental health in the future. Although this latter statement is not specific to mental health, there appears to be a large disparity between those who do use the internet to source information about health, and those who express an interest in doing so. The study described in this chapter explores the reasons for this disparity, to understand what motivates and prevents the use of digital technologies for mental health purposes among older adults. This study aims in part to address the paucity of available research on older adults’ views of technology for mental health purposes.

### 5.1.2 Aims

While previous research has addressed usability requirements of older adults for technology in general, no research has so far explored usability requirements specific to the self-reporting of mood by older adults. This is despite the existence of multiple apps and websites that allow self-reporting of mood on digital devices (Cotton et al., 2014), and the increasing use of such digital devices among older people (Ofcom, 2015). Similarly, older adults’ motivators and barriers to the use of digital technologies to support mental health have not been explored to a great extent (Chapter 3B). The present study aims to address these research gaps, and also to inform data collection methods for subsequent studies.

## 5.2 Research Questions

1. What are the usability requirements of older adults when self-reporting their mood using digital technologies?
2. What are the motivators to older adults’ use of digital technologies to support mental health?
3. What are the barriers to older adults’ use of digital technologies to support mental health?

## 5.3 Approach

This study used a qualitative approach to elicit older adults’ views and experiences. Taylor and colleagues describe the purpose of qualitative work as follows:

“Qualitative researchers develop concepts, insights, and understandings from patterns in the data rather than collecting data to assess preconceived models, hypotheses or theories.” (Taylor et al., 2016: p8)

The present study sought to explore why older adults might be motivated, or not, to engage with digital technologies to support their mental health, and what their usability requirements were. The aim was thus to develop new “concepts, insights and understandings”, as described by Taylor and colleagues (2016, p8).

The design of the data collection aspect of this study was inspired by the COBALT Tools for Engagement (Challenging Obstacles and Barriers to Assistive Living Technologies), which includes the use of interactive activities involving technology during group sessions (Astell et al. 2012; 2016). In this approach, activities are designed to make the users feel comfortable trying out new technologies and presenting their opinions (Astell et al. 2012; 2016). Traditional focus group studies have an emphasis on “asking questions, exchanging anecdotes and commenting on each others’ experiences and points of view” (Kitzinger, 1995). However, focus groups rely on participants recalling anecdotes from memory, perhaps from years ago, and thus subtle details of particular situations may be lost. The COBALT approach allows interaction with technology and other materials during the sessions, meaning that immediate reactions to materials presented can be accurately recorded, since participants are not required to rely on memory.

Kitzinger describes that the downside of group dynamics in focus groups “is that the articulation of group norms may silence individual voices of dissent” (Kitzinger, 1995: p300). COBALT methods go some way to address this concern. For example, participants are encouraged to work in pairs or small groups to try using different forms of technology. All parts of sessions are audio and video recorded, meaning that views participants may only express to one or two of their peers are also captured, reducing the influence of large group dynamics on results. Full details of the methods used are described in Section 5.5.

## 5.4 Pilot study

Before the main study was undertaken, a pilot study was used to establish the feasibility of a subset of the proposed activities, and the feasibility of using different apps and websites with older adults under research conditions. Since many of the apps and websites were recently developed, it was not known if a good sense of their purpose and functioning could be understood and evaluated by participants within the time constraints imposed by the research paradigm. It was important that participants would be able to confidently express an opinion on the tools after using them for up to ten minutes. The pilot study also served as a chance to trial video- and audio-recording devices, and explore the best positioning of these devices.

Seven participants were recruited for the pilot study. This was a convenience sample recruited from among residents of an apartment complex for adults aged 55 and over in Sheffield, UK. The event was advertised through the activity guide that was delivered to all residents. The pilot study was carried out in the communal lounge area of the apartment complex. Audio recorders were placed on the floor between participants’ chairs. Participants and the researcher sat in a circle together, and a video camera was set up behind the chair where the researcher was sitting. The pilot included three novel activities, each with an associated set of materials. While these activities were inspired by the COBALT approach (having elements of interactivity, taking a user-as-expert approach, using small group interaction, using technology interaction within sessions), they were not featured in the original set of activities in the COBALT studies (Astell et al., 2012, 2016). Table 12 gives a description of each activity and describes the materials used in each.

Table 12. Activities used in the pilot study.

|  |  |  |
| --- | --- | --- |
| **Activity** | **Description** | **Materials** |
| (1) Technology icebreaker | Participants discussed what the term ‘technology’ meant to them, and recounted to each other humorous stories of when they had difficulties with technology. They then fed back as whole group, with the researcher taking notes on a whiteboard. | Whiteboard and pens. |
| (2) Activity sorting | Participants worked in pairs to decide whether or not they would use technology for a range of tasks including some related to mental health and some not. They also considered what type of technology they might use to complete these tasks. | Four packs of 10 laminated cards. Each card featured an activity that could be done with or without technology, e.g. send a photo, research a health problem, keep a personal diary. These were written in size 55 text. See Appendix 3 for a full list of the activities and materials used. |
| (3) App interaction | Participants were asked to try out a number of apps and websites and to discuss whether they would use these apps themselves, along with reasons why/why not. Participants tried the apps and discussed in pairs, then fed back to the whole group. | A Samsung Galaxy S3 smartphone with ‘Wellmind’ app installed, a Samsung Galaxy Tab 2 with ‘Five Ways to Wellbeing’ app installed, and a 15 inch touchscreen Eeetop PC with ‘Moodscope’ website open. |

### Reflections

The technology icebreaker (Activity 1) was found to introduce too many different ideas about what technology could comprise for it to be useful. For instance, participants spoke about how the biro pen was a technological invention. This affected participants’ comments in the second task, as when asked if they used technology to keep a personal diary, some said they did because they used a biro. While this may be accurate given a dictionary definition of technology, this did not address the use of digital technologies, which this study sought to explore. The second part of the icebreaker was also problematic as participants found it difficult to think of humorous stories about technology. In light of these findings, the main study was reframed around the idea of ‘digital technologies’ instead of the broader term ‘technology’, and the icebreaker tasks were replaced with a different kind of icebreaker which achieved the same aim of encouraging participants to start thinking about technology and mental health.

The activity sorting task (Activity 2) was successful in generating interesting discussion around technology for tasks related and unrelated to mental health although, as mentioned above, some discussion involved a very broad interpretation of the term ‘technology’ as a result of the icebreaker task. Despite this, it was decided to keep the card sorting activity in the main study, as it had been useful for stimulating discussion relating to the research questions.

The app interaction activity (Activity 3) was beneficial for stimulating discussion on use of apps for monitoring and supporting mental health. However, the ‘Moodscope’ website was deemed inappropriate for a time-limited research session, as reporting mood using this website involved sorting 20 descriptive cards on the screen and this was found to take participants too long to complete. Therefore the app interaction activity was included in the main study, but the ‘Moodscope’ website was replaced with an app entitled ‘Mindshift’.

During the pilot study, a video camera was set up to record participants’ discussions, and audio recorders were used to capture comments while participants were working in groups. On reviewing the video and audio files, it was found that participants had moved their chairs during the study meaning that not all of the participants were in the video frame for the entire session. Positioning of the camera and chairs was therefore reconsidered for the main study. On the audio recorders, it was difficult to hear what some participants were saying when all groups were working at the same time due to differences in the loudness of different participants’ voices. Positioning of the audio recorders was therefore reconsidered in the main study to optimise the recording quality, for instance recorders were placed on tables instead of on the ground so they were closer to participants’ mouths.

## 5.5 Main study

### 5.5.1 Design

The main study took the form of four interactive group sessions with two groups of older adults. Each group attended two sessions of two hours at the University of Sheffield.

### 5.5.2 Participants

In total, 15 participants were recruited for the study: one group of seven participants and one group of eight. Having this number of individuals in each group was considered optimal, since previous work using the COBALT methodology successfully used groups of eight (Astell et al, 2016). Participants were a convenience sample, recruited from two volunteer groups in Sheffield. The opportunity to participate was advertised by the researcher at each volunteer group’s monthly meeting. Both volunteer organisations were aimed at people aged 50 years and older.

### 5.5.3 Inclusion criteria

Inclusion criteria were: aged 50 or over, able to read (with or without glasses), able to hear (with or without hearing aid), no diagnosis of cognitive impairment, and not currently experiencing mental health difficulties. It was made clear that participants did not need to have extensive knowledge of technology to take part.

Here the decision was taken to recruit adults aged 50 and over, for the same reasons as expressed in the inclusion criteria for the second mapping review (see Section 3.6). The development and implementation of the approach described in this thesis may take 15 years or more, by which time people currently aged 50 will be 65, and their experiences using technology throughout their lives are likely to be different to the life experiences of people who are currently aged 65.

People currently experiencing mental health difficulties were excluded from the study because this thesis seeks to develop tools for future prediction - these tools are envisioned to be used with people who are at risk of a mental health condition but who do not, at the time of use, experience any such condition. Thus the study sought the views of people similar to potential users.

### 5.5.4 Ethics and payment

Both the pilot study and the main study were approved by the Research Ethics Committee of the School of Health and Related Research at the University of Sheffield, approval number 003140 (see Appendix 5). Participants received an information sheet and signed a consent form before taking part. They were invited to telephone the researcher with any questions they had before attending the sessions. No payment was provided to the participants, although return taxi fares to the university were offered, and refreshments were provided in all sessions.

### 5.5.5 Procedure

The sessions were organised as a series of facilitated activities, each with a set of materials or technologies. While these were presented and coordinated in line with COBALT principles, the majority of activities were not featured in the original COBALT study – only Show and Tell (Activity 5) was included in the original COBALT study (Astell et al., 2016). Table 13 describes the activities completed by participants. Both groups attended two sessions of two hours each, and completed the activities listed in Table 13 in the same order. Activities were designed to encourage participants to feel confident discussing their own experiences of technology, and to feel confident using and reacting to the different apps presented. In addition to activities where participants worked in small groups of two or three, whole group discussions allowed participants to reflect on their experiences in the sessions. For both groups in the main study, sessions were conducted in the Home Laboratory (Home Lab) in the Centre for Assistive Technology and Connected Healthcare (CATCH) at the University of Sheffield, UK. The Home Lab is a multipurpose space set out as a one-bedroom flat complete with furniture. This setting allows research participants to test new technologies in a home-like environment.

Table 13. Activities undertaken in the main study.

|  |  |  |  |
| --- | --- | --- | --- |
| **Activity** | **Description** | **Materials** | **Purpose** |
| **------ SESSION 1 ------** | | | |
| Icebreaker (1) | Participants introduced themselves and said something that cheers them up when they are feeling down. | N/A | To make participants feel comfortable speaking in a group, and to begin thinking about ideas associated with mood. |
| Activity sorting (2) | Participants worked in pairs to decide whether or not they would use technology for a range of tasks including some related to mental health and some not. They also considered what type of technology they might use to achieve these tasks. | Four packs of 10 laminated cards. Each card featured an activity that could be done with or without technology, e.g. send a photo, research a health problem, and find a solution to a DIY problem. These were written in size 55 text. See Appendix 3 for a full list of the activities used. | To begin to elicit attitudes towards use of technology to support mental health. |
| Vignettes (3) | Participants worked in groups to read and discuss four vignettes. These were paragraphs inviting the reader to imagine that they or their friend were experiencing symptoms of anxiety or depression, though the words ‘anxiety’ and ‘depression’ were not featured on the cards. Participants were asked to consider what they would do if they experienced such symptoms, or what they would advise their friend to do if the card described the reader’s friend to be experiencing the symptoms. | Four packs of four vignette cards were prepared, each describing symptoms of either depression or anxiety without naming either condition. These cards were written in size 14 text. (See Appendix 4). | This activity aimed to engage participants to consider the real-world application of mental health technologies. This was done to contextualise later activities. |
| App interaction (4) | In pairs or small groups, participants tried apps and websites on a smartphone and on tablet computers. Participants were encouraged to try using as many of the apps as time allowed, discuss how useful they were, and decide if they (the participants) would use these apps themselves in their own homes. Experiences were shared as a group. | Apps selected were: Wellmind (Dudley and Walsall NHS Mental Health Partnership), Five Ways to Wellbeing (Somerset Public Health), and Mindshift (AnxietyBC®). Wellmind was presented on a Samsung Galaxy S3 smartphone. Five Ways was presented on a Samsung Galaxy Tab 2 tablet. Mindshift was presented on an Apple iPad Air. | To explore motivators and barriers to use of apps and websites for the purpose of supporting mental health. |
| **--------- SESSION 2 --------** | | | |
| Show and Tell (5) | In an activity inspired by the COBALT study (Astell et al, 2016), participants presented a piece of technology they loved and a piece of technology they had abandoned. They were asked to discuss why they loved or abandoned these technologies. This was conducted as a whole group. | Participants’ own self-bought and self-chosen technologies. | Warm-up, to give participants confidence and feel like experts on the topics at hand (Astell et al., 2016) |
| App interface evaluation (6) | In pairs or small groups, participants evaluated four different ways of self-reporting mood. These were each featured in different apps: (i) Mr Mood, in which users slide full screen emoticons up or down to select the right emoticon to reflect how they are feeling; (ii) Pacifica, which uses a circular slider to reflect different levels of mood; (iii) Five Ways to Wellbeing, which uses five separate sliders to report on different aspects of how users feel; and (iv) NANA mood, a research tool which asks users to press a number between zero and ten to record how much they are currently experiencing each of six different moods (Brown et al., 2016; Astell et al., 2014). Participants recorded comments on feedback sheets which asked them to rate each app on ease of use, intuitiveness, and speed of use. These were used to stimulate thought and discussion among participants. | Apps presented on tablets and a PC: Mr Mood (1Button), Five Ways to Wellbeing (Somerset Public Health), Pacifica (Pacifica Labs Inc.) and a mock-up of the mood reporting module of the NANA homesystem (Novel Assessment of Nutrition and Ageing) (Astell et al, 2014). Mr Mood and Pacifica were presented on Apple iPad Airs, Five Ways was presented on a Samsung Galaxy Tab 2 and NANA Mood was presented on a 15 inch touchscreen Eeetop PC. Participants were provided with A4 scoring sheets to prompt discussion. | To explore usability in mood- reporting. |
| Imagining a future app (7) | Participants were asked to consider different ways an app might respond to low mood scores. After discussing different ideas in pairs or threes, ideas were discussed as a group. All discussions were recorded. | Flipchart. | For participants to consider how data might be used, and to understand how this may affect their motivation to use mood-reporting technology. |

### 5.5.6 Analysis

All sessions were audio and video recorded and transcribed verbatim. The researcher transcribed the audio and video files himself to enable familiarization with the data. All personal identifiable data were removed at the stage of transcription.

Transcripts were analysed using Template Analysis (King, 1998; 2004; 2012). Template Analysis is an approach that aims to “build an understanding of the phenomena of interest” (King, 2004: 267). Applying this form of analysis involved a number of steps, which led to the development of a final template of themes. First, a small number of prior, or “*a priori*” (King, 2012, 2014), themes were developed which formed the basis of the template. King suggests *a priori* themes should “correspond to key concepts or perspectives for the study” (2012 p.430). It was understood that these themes might change or be left out of the final template, despite their initial inclusion (King, 2012). The research questions for this study were concerned with older adults’ views of technology to support mental health. Thus, initial themes for this study related to attitudes to technology and mental health in general, as well as specific attitudes towards technology for mood reporting and supporting mental health (Table 14).

Table 14. *A priori* themes chosen before beginning the coding process.

|  |
| --- |
| Attitudes to technology in general |
| Attitudes to technology for mood reporting and supporting mental health |
| Motivators to use |
| Barriers to use |
| Usability |
| Attitudes towards mental health conditions |

After familiarization with the data, achieved through transcribing audio and video files and reading the transcripts a number of times, the data were coded using Nvivo 11 PC software (QSR International, 2011), in line with the principles of Template Analysis. This was an iterative process which involved reading each transcript and identifying meaningful chunks of text which demonstrated participants’ views. Once identified, these chunks were labelled in one of three ways: either they were assigned directly to one of the *a priori* themes, or they were assigned to new codes which were related to one of the themes, or they were assigned to a ‘floating’ code, where the view represented was not related to any pre-existing theme. New themes were made where two or more floating codes were related. Themes were iteratively renamed, reorganised and removed where appropriate, in line with the approach set out by King (2012). The template was finalised when all transcripts had been read and coded, and all codes had been categorised into relevant themes. The researcher discussed the above process with his supervisory team throughout. The supervisory team had also reviewed some of the data, so discussing the formulation of codes and themes with them served to improve the reliability of the analysis.

## 5.6 Results

The main study was held over four sessions. Group one attended sessions one and two, and group two attended sessions three and four. From group one, six participants attended session one, and seven attended session two. From group two, seven participants attended session three, and six attended session four. Table 15 shows information on participant demographics across the two groups.

Participants across both groups were mainly female (12 out of 15). Ages among all participants ranged from 52 to 88. The first group was older, with a mean age of 68.25 compared to a mean age of 64.14 in the second group. The majority of participants (11 out of 15, 73%) had attended further or higher education. This is a higher proportion than the national average of 32% for 55-64 year olds (OECD, 2014), suggesting that the sample is biased in this respect. The majority of participants (11 out of 15) reported that they regularly used the internet, though this was much higher in the second group (seven out of seven) than in the first (four out of eight). Seven participants reported that they used technology to manage a health condition, although no data are available on what type of technology they use, or what condition they manage. All participants rated their health as fair or good.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Group | Identifier | Age | Gender | Highest education level attended | Technology the participant has experience using | Does the participant regularly use the internet? | Does the participant use technology to manage health conditions? | Participant’s overall health rating (Very poor, poor, fair, good, excellent) |
| 1 | P8 | 66 | F | Further/higher | Mobile phone | No | No | Fair |
| 1 | P9 | 62 | F | Further/higher | Mobile phone, computer, email | Yes | No | Good |
| 1 | P10 | 71 | F | Lower secondary | Mobile phone, computer | No | (no answer) | Fair |
| 1 | P11 | 88 | F | Primary | Mobile phone, digital camera | No | Yes | Fair |
| 1 | P12 | 70 | M | Further/higher | Mobile phone, digital camera | No | No | Fair |
| 1 | P13 | 69 | M | Further/higher | Mobile phone, digital camera, computer, email | Yes | No | Fair |
| 1 | P14 | 52 | F | Further/higher | Mobile phone, digital camera, computer, email | Yes | No | Good |
| 1 | P15 | 68 | F | Further/higher | Mobile phone, digital camera, computer, email | Yes | Yes | Good |
| 2 | P16 | 69 | F | Upper secondary | Mobile phone, digital camera, computer, email | Yes | Yes | Fair |
| 2 | P17 | 68 | F | Upper secondary | Mobile phone, digital camera, computer | Yes | No | Good |
| 2 | P18 | 66 | F | Further/higher | Mobile phone, digital camera, computer, email | Yes | Yes | Fair |
| 2 | P19 | 56 | F | Further/higher | Mobile phone, digital camera, computer, email | Yes | No | Good |
| 2 | P20 | 60 | F | Further/higher | Mobile phone, digital camera, computer, email | Yes | Yes | Good |
| 2 | P21 | 75 | M | Further/higher | Mobile phone, digital camera, computer, email | Yes | Yes | Good |
| 2 | P22 | 55 | F | Further/higher | Mobile phone, digital camera, computer, email | Yes | Yes | Fair |

Table 15. Participant demographics.

### 5.6.1 Template analysis

Table 16 shows the final template after the coding process had been completed. During the coding process, ‘Usability’, ‘Motivators’ and ‘Barriers’ changed position in the template and became higher level themes. *A priori* themes on attitudes to technology in general and attitudes to mental health conditions were eventually removed from the template. A theme on awareness of uses of technology to support mental health emerged during the coding process. Multiple associated subthemes were also generated.

Table 16. Final template including themes and subthemes after coding.

|  |  |
| --- | --- |
| **Themes** | **References** |
| **1. Usability (User-friendliness)** | 70 |
| Big buttons | 3 |
| Big enough to see, ideally without glasses | 10 |
| Cannot rely on transferable knowledge | 5 |
| Like to be able to change answer | 6 |
| Importance of clarity | 26 |
| Importance of portability | 3 |
| **2. Motivators to use of technology to support mental health** | 33 |
| Using technology instead of speaking to a person when feeling down | 6 |
| Dispelling loneliness | 2 |
| Music as therapeutic | 11 |
| NHS sites as a source of information | 4 |
| Distraction/Take your mind off it | 6 |
|  |  |
| **3. Barriers to use of technology to support mental health** | 37 |
| Symptoms of poor mental health may affect readiness to engage | 3 |
| Need for self-awareness and knowledge about mental health to find useful information online | 5 |
| Self-diagnosis using technology is problematic | 9 |
| Technology is inferior to humans | 11 |
| Problems common to technology in general and specifically for technology to support mental health | 9 |
| **4.** **Awareness of technology to support mental health** | 32 |
| Lack of awareness | 6 |
| Meditation and mindfulness | 5 |
| Tools for researching mental health | 4 |

Note: *‘References’ refers to the number of quotes coded to each theme. Coding was completed using NVivo 11 (QSR International).*

Below, themes are explored through the use of example quotes, and the ideas that participants shared are synthesised into research findings.

### Usability

The size of text in many of the apps and websites tested by participants was found to be too small. Participants stated that this would not have been a problem if it had been easy to change the font size within the app they were using.

R: Do you think the tablets are too small?

P17: No.

P21: Well it depends on the font. The only thing I found with the tablets, (P22: Can’t read it can you) The fonts can be small.

P22: That one (pointing to NANA Mood) was nice and big

P21: And for older people they should make them, like, at least 14.

P17 and P18: Yeah.

P22: See when you’re on your own computer at home, you can change your font size. But on these, you have to go into settings, and everything to do it.

Participants remarked on the benefits of a system where glasses would not be necessary to use the technology:

P21: (Discussing NANA Mood) The thing is, everything, Start (referring to the ‘Start’ button within the app), you wouldn’t need glasses, even if you use reading glasses and this sort of thing. (R: Right, ok) That were a big thing with all these, is, eyesight.

The size of the buttons used within the apps was something noticed by other participants as well, suggesting it is not only the visual aspect of applications where size is important, but also the tactile aspects.

P19: You have to have, sort of, your fingertips, don’t you for these little things.

P18, P21, P22: Yeah.

Portability was another important factor for participants when evaluating the usability of different technologies for self-reporting on mood. When asked about screen size, participants in one group commented that they preferred the tablet computers over PCs for their portability, despite tablets having smaller screens.

R: So what do we think about screen size then, you’ve tried tablets, and you’ve tried the bigger screen at the back there, so perhaps for ease of use, which is easier to use?

P12: Tablet.

P9: Mmm, yeah

P15: Cos you can take it with you.

Size and portability represent an area where balance is important – participants appeared to need an adequate size of screen for text to be large and legible and buttons to be easy to press, while at the same time they prefer the device to be portable so they can carry it with them. Some participants preferred larger, portable devices:

P22: That’s why I like my laptop at home, cos the screen’s big.

P18: Yeah, I prefer a laptop.

There were thus a wide range of views on the best combination of portability and screen size, and personal choice was of importance to these participants.

Many of the apps evaluated by participants assumed a certain level of transferable knowledge on the part of their users, for example in the employment of cross symbols to enable a user to close a text entry box, or in the use of a swipe action to move between options. For some participants, these assumptions caused difficulties, as they did not always have the required transferable knowledge.

P19: If you push that cross does that mean it cancels it out? Yes. Right. You do it.

P20: No I think it, does it, oh.

P21: If you’ve never used one before, see I’ve used iPads and computers for a long time so it just come natural to.. scroll on screen, scroll up and down. (unintelligible)

P18: Couldn’t work it out, really.

P17: No.

P9: Right, now ‘apply’, what does that do?

It is clear from these quotes that functions within apps need to be explained more fully for older adult users to feel comfortable using them, and not rely on transferable knowledge, since several participants had to give up on the actions they were attempting to execute.

When self-reporting their mood, both groups of participants mentioned that they would appreciate the ability to go back and change scores they had entered, which was not possible with some of the apps.

P15: The only thing I didn’t like was there was no back button, [..] if you thought oh I answered that question, I’m not sure about if I answered the previous question, maybe I want to change my response, there’s no back button to go back and change my response.

Together, these findings highlight a need for a balance to be struck between simplicity and functionality. There need to be enough functions available to satisfy users’ needs, while keeping the interface from becoming too complex. In addition, participants were confused by technical jargon, demonstrating a need for this to be kept to a minimum.

Usability of technology to support mental health was also affected by the level of instruction provided to participants. Participants struggled to use technology where little or no instruction was given:

P14: Yeah, we just didn’t know what to do. Even though you press…

P9: Not easy to understand at all.

P14: Nothing was coming out, so we’re just like, stuck.

Instruction within apps, printed instruction manuals and verbal instructions provided by another person were all discussed within the sessions. Participants commented that they liked and sought instructions for new technologies, but complained about the small size of text in many instruction manuals.

P17: I need instructions!

P15: I was thinking the instruction manual was very small. The little tiny bit of paper they send you.

Many participants stated that their preferred way of learning how to use technology was to be shown how to use it by another person:

P20: I think if we’d had someone explain how to use it, I would have found it really good to use, but it was just trying to manoeuvre our way round it.

P15: I think the problem there is, if it’s not well explained, it’s not going to be quick! You’re trying to find out how to do it, which takes extra time, whereas if somebody gave you an explanation in the first place, it might be quick once you know what you’re doing.

When it came to healthcare technologies, one participant made clear that she would expect training to be given by the healthcare provider.

P15: I suppose if I think I was given anything by anybody in authority at hospital or anything, I would expect them to explain to me how to use it.

These findings serve as a reminder that healthcare technologies always form part of a larger network of people and pathways that make up a service. They suggest that any service involving technology for monitoring mood should provide clear instructions and/or personal training to its users.

### Motivators

Participants described many reasons they may be motivated to use technology to support their mental health. Firstly, participants described feeling reticent to discuss feeling down with friends or family members for fear of spreading misery.

P9: You could ring a friend I suppose, as well. I usually don’t have any.

P18: I’d be worried about passing misery on, you know what I mean. (P17: I know) yeah.

P18: Don’t want to pass it on.

P17: No

P18: That’s the thing.

P20: It depends if they’re fed up of you though, doesn’t it (laughs), (P22: Eh?) If your family have heard it again and again and they think oh, here we go again.

Using technology in place of speaking to a person was seen as a potential way to prevent ‘bothering’ others:

P22: Well I think if your, if your computer was set up that way, if you spoke to it and said I’m in a low mood today, what should I do? Cos it’s like, you don’t have to bother anybody then because the computer’s programmed to s.. well like one of them android things, just to say, well I suggest you go for a walk or, whatever it churns out, really. I think that’s good. Cos you don’t always want to bother somebody, like these ladies have said, you don’t.

Loneliness was a particular source of concern for participants. Many recognised that loneliness was detrimental to health. In the first quote below, participants discussed having family members living close by.

P9: And if you’ve got that then you’re very very lucky because you know, as you must know (P13 name), you know with these people that you deal with, they don’t have anybody to, and nobody close to them or anything, you know, anybody to talk to.

P13: I work with too many that’s… lonely.

P9: Don’t have anybody, yeah.

P11: It is a killer, isolation is a killer, and to get these people out, you’ve got to work very hard at it.

Participants discussed how they and others they knew used technology to alleviate loneliness:

P17: Like people can’t get out and they haven’t got technology, like my mother in law, just has the telly on all day.

[..]

P22: I understand what she’s going through, cos when I was ill, I had to have visual things, (P20: Yeh) I will switch the radio on sometimes but the visual contact of the TV makes me feel that I’m not on my own. (P20: Yeh) And you don’t have to watch drivel, you can watch the Victorians, and how they survived and things like that you see.

P17: But she doesn’t watch anything like that, she watches like come dine with me and…

P22: It’s company, you see people.

However, one participant made clear that using technology for this purpose should be a last resort rather than the norm:

R: Mm, so what about talking to a computer then, do you think that is any kind of substitute for a person? What do you all think?

P21: Well, it’s last resort, like I say, if you are, on your own, then yeah, it…

P20: It’s better than nothing, isn’t it.

Participants mentioned that they were motivated to use music-playing technology because it helped to change how they were feeling. Participants used a wide variety of different technologies to source music, including CDs, tapes and vinyl, as well as online sources like Youtube.

P13: Relax and unwind. You would not use digi-, well, you could use digi-, er, technology for that because you could play records, (P12: Yeh, play CDs) play CDs, watch telly.

P12: Yep, yep, I find technology there very important.

P21: If I want a lift, an uplift I go and put rock and roll on, or summat a bit lively.

P17: I put music on. Put Youtube on and type in what I fancy listening to.

P16: Yeah, I like the Youtube for all sorts. Er but music is, does relax me.

These quotes show that older adults view music as an effective way to alter their mood. Video games were also seen as helpful, as a way to distance oneself from one’s problems, improving present mental state through distraction:

P19: (Discussing the game feature of the app Wellmind) Oh there’s playing a game with s.. a, erm, a kind of a snake and an apple, you have to move a bar and hit the apple and things and..

P18: Which I’m not sure what that’s got to do with it, but..

P19: It’s er,

R: It’s a copy of that old snake game you used to get on old mobile phones.

P19: It’s a sort of erm, I suppose it’s a, a sort of a concentration of the mind or something, yeah.

P21: well it er, you stop thinking about something cos it just…

P20: Yeah you’re focusing on something

P22: Yeah

P21: It’s if like, sommat’s happened or you know, so it will take your mind if you’re, concentrating on one thing. Just gives you that bit of respite.

When it came to using the internet to research mental health conditions, one participant felt confident using the NHS website and saw the potential benefits of using it. Here, she reacts to a vignette in which the person is imagined to be experiencing symptoms of depression:

P20: I suppose for this one, this red one, you could like, look up a.. NHS site, Moodzone, they do audio tapes that, they do tapes that actually give you some inspiration towards feeling better.

In summary, the quotes in this section demonstrate how participants saw benefits to the use of internet and digital technologies for their mental health, and were motivated to use such technology to change their mental states.

### Barriers

Despite discussing a number of motivators to the use of technology to support mental health, participants also recognised a number of barriers to its use. First, participants suggested that symptoms of poor mental health were likely to affect readiness to engage with technology.

P8: And if you’ve gone home, and you’ve had a stroke and you’re depressed, you’re going to get even more depressed (P9: mmm, P15: Yes) because you can’t understand the gadget (P9: Yeah, P10: Yeah).

P15: Yes, that’s another point, if you’re older and you’re depressed or you’ve got depression, you’re going to need even more instructions, not less, because you’re not going to feel like settling down to...

P18: I do wonder if you could be so low, that you couldn’t be bothered to go on a computer.

Some participants mentioned that fear may also prevent them from engaging with technology for self-reporting mood. This was evident when they considered what a computer might suggest an older person should do when they had reported consistently low mood.

R: Go and see a doctor?

P17: Yep.

R: OK. So you think that would be a good thing for the computer to tell someone to do?

P17 and P18: Yep.

P20: (shakes head) I’d be scared.

R: You’d be scared? Scared, can I ask scared, why you’d be scared?

P20: Cos I’d think that the computer had assessed me, and that I was… (P17 and P18 laugh) (P17: Mad!) desperate! (laugh) I don’t know.

P22: That there’s a little man in there, looking at you!

P20: I don’t know, yeah.

This quote demonstrates one participant’s fear around use of technology for automatic assessment of mental health. Participants discussed how such fear might manifest in themselves and other older adults, and the possible consequences this could have.

P15: where the fear is of something unpleasant happening to you, like you’re being taken, taken into an institution, then you’re not going to admit any symptoms that you might think could be interpreted that way, so all that’s going to happen is that yes, you can get all the technology in the world to put people’s moods in, they’re not gonna admit if they feel depressed.

This quote demonstrates how fear around assessment of mood might come from users fearing what will happen to them if a problem is detected. As well as these concerns, there was a sense that human contact was superior to technology when it came to dealing with mental health, and that dealing with people was preferable to using technology.

P20: If you really feel down and you really need some help, can a computer… (P18: Yeah) don’t you think physical contact or speaking to someone is far more important?

P19: Yeah, I think a computer doesn’t, assess, assess the person, like a person might.

P20: Yeah

P18: Yeah cos you can visually assess a person can’t you, rather than relying on what they’re saying, you can look at someone and..

P20: Yeah

P19: And also the feedback, it’s the personal contact, that might be better

P9: But when you’re talking about health, a doctor needs to look at you, really, and feel you, you know, you can’t do it over a phone.

P20: I think you need a person.

P18; Yeah

P20: You need somebody to, probably, bring you out of it.

However, these views contrasted with views mentioned earlier around the benefits of technology when an individual did not feel confident speaking with a person. These differing views on the best way to support mental health indicate the benefits of providing a choice to older adult users, to use technology or not.

Many barriers to older adults’ use of technology in general, described elsewhere in the literature, were also found to be problematic for participants in this study, when considering the use of technology for the specific aim of supporting mental health. These included general fear of using technology (“P20: I don’t know why I’m afraid of downloading apps”), difficulties reading text (“P22: The print’s too small, can you make it bigger?”), and a sense that using technology is time-consuming (“P9: When you go on a computer, you’ve got to log on and all that bah blah blah”).

### Awareness of technology to support mental health

Participants demonstrated an awareness of multiple uses of digital technologies to support mental health. These included online doctors’ appointment booking systems, NHS and other websites for researching health conditions, meditation apps and Youtube videos for relaxation. Many participants made reference to CDs, tapes, and online resources for mindfulness, though one participant was unsure about the way in which the online resources would work:

P8: How does mindfulness work online?

Another participant found it difficult to use such resources alone:

P22: I think with this it all depends on the individual, cos we’re all different. (P18: That’s true) You might find some groups that really like this and some that don’t. I’ve got those tapes that they’ve given me to teach me how to relax. I find that, I can’t do it on me own. But I can do it say if you was with me, and we just, meditated together, I could do that. But on my own, I’m finding, ‘oh I haven’t washed the pots, I haven’t done this, I haven’t done that, you know.

Thus, although participants showed an awareness of many ways in which digital technologies can be used to support mental health, awareness alone is unlikely to be enough for participants to start using such tools successfully.

## 5.7 Discussion

### 5.7.1 Principal Results

The aim of this study was to use a series of interactive activities to understand the usability requirements, the motivators, and the barriers to older adults’ use of technology to report mood and support mental health. Below, findings from the study are discussed in relation to existing research on older adults’ use of technology, including that for mental health purposes. Findings are considered within the scope of the thesis, that is, relating them to the development of a tool for predicting depression and anxiety. The research questions for the study are used to structure this section.

1. What are the usability requirements of older adults when self-reporting their mood using digital devices?

This study has found that older adults require text and buttons to be large, or ideally adjustable in size, so that they can be used without causing eye-strain. This reflects research by Chen and Chan (2013) which showed that older adults appreciate large fonts when using technology, and that older adults fear eye-strain through using technology.

Using large fonts and buttons necessitates the use of a screen size large enough to accommodate these. However, participants also spoke of a preference for devices to be portable. Mitzner et al. (2010) reported that with regard to technology in general, older adults found the portability of devices to be of importance. The findings presented here suggest that while this is the case, a trade-off is required between the portability of the device and the size of the screen for presenting text and buttons which are legible and can be easily activated. Vaportzis et al’s focus group study examining barriers to older adults’ use of touchscreens also reported disagreement between participants on the best size of device owing to differing opinions on the importance of both portability and screen size (2017). In the present study, participants’ comments suggested that mobile phone screens may be too small for older adults to comfortably report on their mood. With their larger screens, tablet computers appear more appropriate, although some participants expressed a preference for using a laptop or desktop computer. As such, allowing older adults to use their own preferred devices to report on their mood may be most motivating of continued use.

Many participants in this study experienced difficulty understanding jargon and practices commonly used in apps, for example the use of cross symbols and ‘apply’ buttons. Similar findings have been reported in (Eisma et al., 2004), where scrollbars are used as an example of an object with which older people may not be familiar, highlighting a cultural difference between expert users and the majority of older adults. Ensuring that function buttons are clearly explained, while avoiding the use of technical jargon, is likely to ensure that older adults feel comfortable and confident using mood-reporting technologies.

Older adults in this study preferred having the ability to go back to change or check answers to previous mood questions. Several participants remarked on the lack of such a function in some of the apps they tried. While previous research has highlighted the importance of simplicity in technology destined for use by older adults (Chen & Chan, 2013), results here suggest that a trade-off is to be made between simplicity and functionality in the development of such technologies.

1. What are the motivators to older adults’ use of digital technologies to support mental health?

This work has highlighted several reasons that older adults may be motivated to engage with technologies to support their mental health. Firstly, participants expressed being motivated to use technology to alleviate low mood, for example by listening to music, playing games, or watching television. Prior research has highlighted that technology can provide older adults with opportunities for enjoyment and fun (Astell, 2013), and participants here demonstrated that this can have the added benefit of alleviating low mood.

Secondly, some participants believed that technologies including television and radio could help to ease feelings of loneliness, and this is also supported by prior research (Cotten et al., 2013). Alleviating loneliness was thus another motivator to engage with technology. The use of technology for these purposes reflects how older adults seek to manage symptoms of low mood by themselves, without reaching out to friends or family. Further comments made by participants suggested that this can be caused by a fear of becoming a burden. Peek et al (2015) also found that older adults were keen to avoid being a burden and used technology to this end. The will of older adults to be self-reliant can thus be seen as a motivator for them to engage with technology that helps them to manage their mood and feelings of loneliness.

In addition, the finding that some older adults are comfortable using music and video games to change how they feel could have implications for the design of a mood-reporting app. For example, if scores reported by a user predict later depression, the app could suggest appropriate music to improve their mood. If participants’ scores are found to be predictive of anxiety, distracting activities like playing simple video games, or relaxing activities like watching TV, could be recommended. Further research would be required to validate these approaches, and consideration should be given to the lack of human intervention in such approaches.

1. What are the barriers to older adults’ use of digital technologies to support mental health?

The main barrier perceived by these participants in relation to mood-reporting was fear. This seemed to originate from participants’ belief that something bad would happen to them if the mood-reporting software detected a problem with their mental health. Participants expressed that this fear could lead to users falsely reporting more positive mood states than they truly experienced in order to avoid potential consequences. These findings reflect the fact that stigma can affect help-seeking for mental health issues among older adults, as found by Préville et al. (2015). However, the study here has gone further by exploring what older adults imagine the actual consequences of predicting future mental health conditions might be (being “taken into an institution” - P15), thus illustrating an example of thought patterns that underlie this phenomenon. This finding could inform preparatory documents and explanations that might be provided alongside mood-reporting software to address users’ fears. For example, these materials might explain that if a prediction of poor mental health were made, further assessments might be carried out, but reporting using the technology itself would not lead to sectioning under the Mental Health Act.

Some participants felt that, if affected by a mental health condition, they would prefer to speak with a human health professional about their experience rather than use technology. Preference for human management of healthcare and lack of trust in technology have also been found to concern older adults in prior research (Fischer et al., 2014). These findings can inform the way in which predictive technologies should be presented to older adults - users should be reminded that the technology will not replace care by human health care professionals, but will merely assist them to identify those with greatest need. More broadly, these findings suggest that the exclusive use of technology-based solutions to treat mental health conditions may not be suitable for an older adult population at present.

Participants also discussed how symptoms of poor mental health may affect readiness to engage with technology. Comments included the fact that feeling depressed was likely to magnify difficulties with usability of digital technologies. As discussed in Section 5.5.3, the target users of the approach discussed in this thesis are those who are known to be at risk but have not yet been identified as having a mental health condition. It could be argued therefore that since the approach seeks to detect early signs of depression and anxiety, it is unlikely that poor mental health would affect the use of the mood reporting approach discussed here. However, the models presented in the previous chapter seek to detect feelings of tiredness and sadness, which are symptoms of poor mental health which may affect readiness to engage. Further research would be required to understand this phenomenon more fully, for example what severity of depression or anxiety causes older people to disengage with technology?

### 5.7.2 Contribution to knowledge

While there is a significant research literature available on older adults’ use or non-use of technologies for different purposes such as communication, entertainment, and supporting health (e.g. Chen & Chan 2013; Peek et al. 2015; Greenhalgh et al. 2013; Mitzner et al. 2010), and also on older adults’ attitudes to mental health conditions (e.g. Smyer & Qualls 1999; Préville et al. 2015), very little research has examined older adults’ attitudes to the use of technology for monitoring mood and supporting mental health (see Chapter 3B). As the number of older adults using digital technologies increases (Ofcom, 2015), and with the importance of prevention in healthcare becoming more recognised (Windle, 2015) self-management using technology will become more important over the coming years. Indeed, the NHS Five Year Forward View for Mental Health strategy has highlighted the importance of making use of technology in mental health care (Mental Health Taskforce, 2016). By exploring issues around older adults’ use of digital technologies to support their mental health, this study has contributed knowledge that could be useful not only for other studies within this thesis, but also for developers working on applications which feature an element of mood reporting or supporting mental health for older adults.

New findings presented in this work include that older adults have differing preferences over the type of digital technology they feel comfortable using. Therefore, allowing them to use a device of their preference to report on their mood is likely to be motivating of continued use. The study has also shown that self-reliance can be a motivator to the use of technology for older adults. It demonstrates too a need for a balance between simplicity and functionality in mood-reporting technologies. It has also illuminated possible reasons older adults might fear predictions of mental illness, allowing these fears to be addressed in new technologies. Further work is needed to better understand how recommendations such as the use of music and games might alleviate symptoms of depression and anxiety in older adults, and also to understand how mental distress and mental illness might affect older adults’ readiness to engage with technology.

### 5.7.3 Strengths and limitations

With a convenience sample of just 15 participants in the main study, the sample size used here is small. Participants were unrepresentative of the population in that they had a higher than average level of education. The possible implications of this on the generalisability of findings is not clear - while some research has found a correlation between higher levels of education and greater use of technology among older adults (e.g. Vorrink et al., 2016), other research has failed to find such an association (Scanlon et al., 2015). Demographics data collected in the present study indicated that participants had a variety of levels of experience with technology, including some participants with relatively little experience of using technology, meaning views were represented from those with and without extensive knowledge of technology.

The main sessions for this study were held in a technology laboratory at a university, and thus the environment may have encouraged participants to provide responses which were more positive about technology than they truly felt. However, the use of the COBALT approach was purposefully chosen to engender confidence in participants to express their views honestly and openly (Astell et al., 2016), and the range of both positive and critical comments given by participants on the technologies presented suggests that they were not unduly influenced by their environment.

Another limitation is that the study only included a limited number of apps owing to time pressures, meaning it may have failed to offer participants mood-reporting solutions which were preferable to those available. However, the selection of apps to be used was carefully considered so that a broad range of possible mood-reporting methodologies could be explored, and participants were able to express a range of views on these methodologies which contributed to the findings.

### 5.7.4 Summary

This chapter has presented key issues in older adults’ motivation to use digital technologies to report on their mood and support mental health. These issues have been discussed in the context of developing an algorithmic approach to the prediction of depression and anxiety in older adults. The next chapter presents a study in which data were collected from older adults to validate the algorithms for predicting depression developed in Study 1, and to explore the potential for developing an algorithm for the prediction of anxiety. The data collection methods described in the next chapter are informed by the findings of the study presented in this chapter. As such, the findings presented in this chapter are crucial to the approach taken in this PhD, as well as making a contribution to knowledge more generally.

# Chapter 6 – Collection of primary data to validate depression models.

## 6.1 Introduction

This chapter describes the collection of primary data on older adults’ mood and appetite, along with later depression and anxiety measures. These data were collected for two purposes. The first purpose was to prospectively evaluate the models developed in Study 1, and this work is presented in Chapter 6. The second purpose was to develop a model for the prediction of future anxiety. The description of analysis methods, the results and the discussion relevant to this purpose are discussed in Chapter 7.

### 6.1.1 Rationale

As described in Chapter 4, Study 1 explored the application of a number of machine learning techniques to self-reported mood and appetite data from older adults, in order to develop models to predict future depression status, as measured by the GDS. The final approach used in that study was the application of the LASSO with logistic regression. The approach involved cross-validation, using the same sample to both train and test the two models developed. An ROC analysis was used to estimate the predictive ability of each model. The resulting areas under the ROC curves were higher than those reported in many other studies where machine learning had been applied to the future prediction of depression. However, the sample used in Study 1 was relatively small, and the analysis of the data was retrospective – the data were not initially collected for the purpose to which they were put. Study 3, presented in this chapter, describes a prospective study which was conducted to validate the models developed in Study 1.

### 6.1.2 Research question

The research question to be explored is as follows:

* Do the models developed in Study 1 function prospectively to predict the future occurrence of depression in a similar sample of older adults?

### 6.1.3 Study design

Prospective validation of a model involves collecting data for the explicit purpose of applying the model to new data. This experiment therefore required older adults to report on their sadness, tiredness and relaxation once every day over a period of one week, then to complete a GDS test after 9 weeks. Collecting data on these variables would be sufficient to prospectively validate the models.

### Additional measures

One aim of the thesis was to explore the potential for developing a model to predict future anxiety. Therefore, the decision was taken to collect additional data in Study 3. The decision was taken to collect all six of the mood measures included in the Study 1 analysis in the present study, since it was thought that they may reflect symptoms of anxiety. Furthermore, the HADS anxiety subscale was included in the study so it could be used as the outcome variable for a new approach to explore the prediction of anxiety. The HADS anxiety subscale was chosen from among anxiety scales because it had more questions than the GAD7 but fewer than the GAI. This allowed a relatively detailed scoring of participants’ level of anxiety without being so burdensome as to discourage participation. Reducing burden was seen as important, as participants were offered no incentive to take part in the research.

## 6.2 Method

### 6.2.1 Participants

The target population for this study was older adults aged 65 and over. This was the same age range as those in the NANA validation study (Chapter 1). The decision was taken to keep the age range the same as that on which the original models were developed, to reduce any possible bias from age-related differences in mood patterns. Full inclusion criteria were as follows:

* Adult aged 65 or over
* Able to use email and a web browser on a computer, tablet computer, or mobile phone.
* Available to respond to emails every day for one week.

Exclusion criteria were as follows:

* Currently experiencing severe mental health difficulties, since the study sought to predict future onset.

Recruiting participants who were regular users of email ensured that they had access to a form of digital technology, and that they were comfortable using the technology they chose. Findings from Study 2 (Chapter 5) indicated that allowing a choice of digital technology in line with preference is likely to be motivating of continued engagement.

### 6.2.2 Sample size

A sample size calculation was undertaken to determine the number of participants necessary for an ROC analysis. The sample size calculation formula came from (Obuchowski, 2005), and required the following input variables:

* Intended percentage power
* Area under the curve
* Type 1 error
* Allocation ratio (*k*)

For this study, the intended percentage power was set at 80%. In Study 1, the area under the curve reported for the model with a cut-off of five was 0.80, and this value was used in the calculation. The type 1 error was set at 5%. The allocation ratio (*k*) describes a relationship between the number of true positive cases and the number of true negative cases in the outcome measure of interest in the sample. Since this was a prospective study, these figures were unknown at the beginning of the study. Therefore, this figure was estimated based on prevalence of the condition, as per the recommendation in (Obuchowski, 2005). The following formula is recommended for this purpose in Obuchowski’s paper:

*k* = (1 – PREVp)

PREVp

where PREVp is the prevalence of the condition in the population of interest and *k* is the allocation ratio. The prevalence of depression reported in the literature varies considerably depending on the setting [0.9 – 42% according to (Djernes, 2006)]. Given this variability, the prevalence for this calculation was estimated based on the NANA validation study sample, since recruitment methods for both studies were similar, so a similar rate of depression was expected. The number of positive cases of depression (participants scoring above a cut-off of five on the GDS) at the end of the NANA validation study was 7 out of 40. As such, PREVp was set at 7/40 = 0.175, giving *k* = 4.7. The allocation ratio was therefore set at 4.7.

The sample size calculation was made using EASYROC (http://www.biosoft.hacettepe.edu.tr/easyROC/), an online tool which computes the required sample size using Obuchowski’s method. When the above values were entered for computation, the required sample size was calculated to be 35 participants.

### 6.2.3 Recruitment

Similar to the NANA validation study, participants in this study were recruited through the University of the Third Age (U3A) and through a database of older adults managed at the University of Sheffield. In contrast to the NANA study however, participants were not recruited through local newspapers, in order to eliminate the costs associated with this form of recruitment. The U3A is an organisation with divisions present in many regions of the UK. Many divisions have websites with contact details for local organising members. An introductory email was sent to provide some information about the study and asked if these organising members would forward the email to their membership, or alternatively advertise the study with a poster at their next members’ event. Recruiting online and at distance allowed for greater reach of the study, increasing the number of potentially interested parties.

Given the time required of each participant, a high drop-out rate was expected from the study. To mitigate this, the recruitment strategy aimed to over recruit. Initially, 20 divisions of the U3A were contacted to seek help with recruitment. Areas chosen were those local to the Sheffield area and those that other researchers had been able to source participants from in previous research projects. An estimate was made that for each division contacted, an average of three participants may be recruited, giving a sample of 60 participants. This estimate was based on the experiences of another PhD student who had used a similar recruitment method. However, after contacting twenty divisions of the U3A and allowing a period of several weeks for them to respond, it became clear that many divisions would not be able to supply the three participants envisaged. Reasons for this included managerial arrangements around promoting research studies, for example members from four divisions responded that they would have to discuss the project at the next management meeting before approaching their membership, and this would have taken too long for recruitment to be undertaken within the allotted time. Two other divisions refused to promote the study because they had already promoted many university studies recently. Due to these difficulties, along with low interest from members of the divisions who had promoted the study, the decision was taken to contact more divisions of the U3A. Another 18 divisions were contacted, taking the total to 38.

Interested members were asked in the email and on the poster to email or phone the researcher to find out more. Those who did were sent a link to a full information sheet online and were asked to give informed consent by clicking through an online consent form. The information sheet and consent form were presented via Survey Monkey, with a unique identifying keyword provided to each participant to pseudonymise the data.

Survey Monkey holds all data collected on its platform in the United States, and therefore ethical use of its services in research depend upon alignment with Principle 8 of the Data Protection Act 1998, which requires that data held outside the European Economic Area are protected by an “adequate level of protection”. While Survey Monkey is covered by an EU – US privacy agreement, its statement with regard to this agreement declares that data may be collected, stored and processed for a number of different reasons. As such, the exact use Survey Monkey may make of any data stored on their servers is unclear. Thus, in the interest of avoiding potential infringement of data protection laws, pseudonymisation was used to protect participants.

### 6.2.4 Ethics

This study was awarded ethical approval from the Research Ethics Committee of the School of Health and Related Research at the University of Sheffield, approval number 012408 (see Appendix 6).

### 6.2.5 Pilot study

In Study 2, there was no consensus around the most preferred device to use to report on mood. Participants had varying preferences based on what they found most easy to use, what they were used to, and what they had used in the past. In the present study, participants were required to use digital technologies to report on their mood and respond to questionnaires. A solution was sought which would allow participants to use a device of their preference, while keeping the presentation of questionnaires and the daily mood report consistent for all participants. It was not possible to develop a cross-platform software solution owing to a lack of funding and time for development. Instead, a web-based approach was used to facilitate cross-platform accessibility.

The use of Google Forms was initially considered appropriate given its simple presentation style, so versions of the daily mood report on Google Forms were piloted with two older adults using iPads. The pilot participants commented that the text on Google Forms was too small to read, and that the ‘Send feedback’ button (present on all Google forms) was distracting and confusing. Since font size could not be changed using Google Forms and the ‘Send feedback’ button could not be removed, Survey Monkey was chosen for use in the study instead, since it did not feature such a button, and with a Gold membership subscription, it allowed for font sizes to be changed.

### 6.2.6 Procedure

The procedure for this prospective study was based on the original NANA validation study. However, the procedure was adapted in several ways.

The NANA validation study involved the distribution and collection of 15-inch touchscreen computers, on which the data collection software was installed. That study was undertaken by a large team of researchers, based across two UK cities. This meant many people were available to distribute, maintain and collect devices across the two cities. In contrast, the present study was conducted by one PhD student based in one city. This meant fewer human resources were available to distribute and maintain devices. In addition, the limit to one city caused there to be a reduced population of potential participants from which to recruit. Thus, two key issues needed to be addressed in the present study:

(1) *Sample sufficiency*. To ensure sufficient recruitment in line with the sample size calculation, the decision was taken to use online methods to recruit for the study across several counties in the UK, thereby expanding the population of potential participants.

(2) *Resource management*. The study inclusion criteria stated that participants should have access to a digital device (laptop/desktop computer, iPad, other tablet computer, iPhone or other smartphone) and be regular users of email. While this may have biased the sample toward a demographic with a higher income and/or educational attainment, this was considered a worthwhile trade-off against the reduced requirement for provision of costly equipment which would have to be distributed and maintained.

In addition to the above practical considerations, allowing participants to use devices with which they were already familiar was expected to be beneficial according to the results from Study 2. Accordingly, in the present study, participants were allowed to use any device on which they felt most comfortable to self-report the required data.

#### Demographics questionnaire

Once participants had given their consent, they were emailed a link to click through to a demographics questionnaire. This questionnaire was provided using Survey Monkey. The questionnaire asked for participants’ unique memorable word to pseudonymise their responses and protect their personal information. Participants were asked to report their age range, whether they had one or more diagnosed medical conditions, whether they took one or more prescription medications, smoking status, difficulty with activities of daily living, experience with different forms of technology and whether they were regular users of the internet. These data were also recorded in the NANA validation trial, so collecting them here allowed a comparison to be made between the samples in Study 1 and Study 3. Following these questions, participants answered the 15 item GDS and all questions on the HADS, using Survey Monkey, to produce baseline measures.

#### Daily mood report

The researcher monitored the completion of demographics questionnaires and baseline depression and anxiety scales every morning of the study. When a participant had completed the full questionnaire, the following morning a link was sent to them, along with a reminder of their memorable word, so they could complete the daily mood report for the first time. The daily mood report asked participants to report how happy, sad, tired, alert, relaxed and hungry they were on a scale of 0 to 10, with the added option of ‘No answer’, as per the NANA validation study (Astell et al., 2014). The question and answer text were presented in font size 16. Participants then received a similar email every morning for the next six days, for them to report on their mood. These emails were sent regardless of whether the participant had completed the previous day’s report, in order to replicate the system of prompts used in the NANA homesystem, which also alerted users to report on their mood every day, irrespective of the previous day’s activity. In the present study, participants were instructed to complete the questionnaire at whatever time they would usually check their emails during the day.

At the end of the seven days of mood reporting, participants were reminded that there would be a gap of nine weeks before the next contact from the researcher. After this period had elapsed, each participant was sent a link to a final questionnaire which included the 15 item GDS, the HADS and a question asking them to indicate (using a ‘tick all that apply’ format) which device(s) they had used to report their mood during the study.

### 6.2.6 Data analysis

#### Data cleaning

While the intention was for participants to respond to emails every day to report on their mood, inevitably on certain days, participants were unable to complete their mood report. However, a minimum of three days’ mood reports were deemed necessary to reduce the risk of bias from daily fluctuations in mood, as per Study 1 (See section 4.5.1). Therefore, data from participants who reported their mood on fewer than three days were removed from the dataset.

Some participants attempted to ‘catch up’ on days they had missed by completing two or more days’ mood reports on the same day. This was problematic, as the models assumed that each mood report was scored on a separate day. As such, the data from the current study were cleaned in the following ways.

If a participant had reported on their mood twice on the same day and they had reported the same values in both reports, one of the entries was deleted and one was kept, as it was clear that participants were only reporting how they felt on one particular day. However, if multiple mood ratings on the same day were given different scores by a participant, both entries were deleted as it was not possible to know which of the two reflected the participant’s mood on the day in question (i.e., the participant may have tried to think back to how they felt on a previous day, which was unlikely to be accurate).

In addition, data were excluded when participants had not correctly provided their memorable word, for example if they had left the field blank, or if they had given an incorrect memorable word. In these cases, it was not possible to identify who the data had come from, and these entries could not be analysed. Furthermore, where participants had recorded more than seven entries (by following the link in one email more than once), any data beyond the seven required days were excluded from analysis.

The Survey Monkey output data file included details of the date and time when participants started and finished each mood report. These showed that some participants had begun mood reports on one day and finished these up to a week later. Any such entries (where start and completion dates fell on different days) were also excluded from the analysis.

In the answers to the HADS questions in the follow-up questionnaire, some answers were left blank by some participants. To handle this, the subscale half mean approach advocated by Bell and colleagues (2016) was used: For each question where no answer was provided, the mean item score was used for that question.

#### Analysis procedure

Initially, an exploration of the data was made using descriptive statistics. This included exploring participant demographics to compare them against the participants in Study 1, as well as visualising overall trends in mood reporting and GDS results. Data on the devices participants had used to complete the study were also explored.

The aim of the present study was to evaluate the performance of the two models developed in Study 1 in a separate group of similar older adults. Study 1 applied logistic regression with a shrinkage and selection operator to mood and appetite data collected from 40 older adults to predict depression status according to the GDS nine weeks later. This process included the use of cross validation to avoid overspecialisation. Two models were developed using this technique, one using a cut-off of five points on the GDS, and one using a cut-off of three. In both of these cases, the selection operator selected two input variables in the prediction model. Both models were evaluated using a ROC curve analysis.

To test these models prospectively, data from the present study were first aggregated per participant by taking the mean of each mood variable across the days they reported their mood. For each participant, six values were produced, one for each of the mood variables.

The two models from Study 1 were then applied to these data to generate predicted probabilities of each participant scoring above the cut-offs of three and five points on the GDS at follow-up. These predicted probabilities and their corresponding follow-up GDS statuses were then used to produce an ROC curve for each cut-off, and values for the area under the curve were calculated in each case.

## 6.3 Results

### 6.3.1 Participants

Sixty-three people contacted the researcher about the study. Ten of these were excluded for being younger than the minimum age of 65 stated in the inclusion criteria. Another eleven people failed to respond when they were sent the information sheet and consent form. Three individuals completed the baseline demographics questionnaire including the GDS and the HADS, but were found to score in the extreme range for depression and/or anxiety on the GDS and HADS. To ignore this would have been unethical, so these individuals were sent an email to advise them that their scores were indicative of ‘low mood’ or ‘anxious feelings’, with a recommendation that they visit their GP if concerned about this. Since these individuals may have sought medical advice, it was decided to exclude them from the study, because there was a chance they would visit their GP and be provided with pharmaceutical or psychological treatment, which could have biased their results on the mood reports or follow-up GDS and HADS tests.

One person withdrew from the study while completing the GDS, as they found it difficult to answer the questions in the GDS with a simple yes or no. Two participants failed to complete the follow-up questionnaire, and their data were also therefore excluded from the analysis.

After data cleaning, 36 sets of data were included in the analysis, each of these including six or seven daily mood reports along with demographics questionnaires and GDS and HADS anxiety results at both baseline and follow-up. This figure was greater than the figure of 35 required for 80% power according to the sample size calculation.

Table 17 shows demographics data for participants in Study 1 and Study 3. Data for Study 1 were collected in 2011, while data for Study 3 were collected in 2017. The two studies feature participant groups with similar mean ages, although those in Study 3 were slightly younger. There was a greater proportion of female participants in Study 3. There were fewer participants with one or more medical condition in Study 3, but this was a small difference, and the marginally younger sample may have been the reason for this. Rates of difficulties with activities of daily living were similar across the two studies for most questions, although fewer participants in Study 3 reported difficulties shopping for food. The number of participants reporting prior experience using a mobile phone is similar, but slightly greater in the more recent cohort. This reflects an increased uptake of mobile phones in this age group more generally (Ofcom, 2015). There was also an increase in the proportion of study participants reporting experience using a computer, and also in those being internet users. These increases may reflect the requirement in the present study for participants to be frequent users of email, though some of the increase may also be explained by the increase in use of email and digital devices in the target age group over the past six years (Ofcom, 2015).

Table 17. Participant characteristics table for the final participants (after data cleaning), compared to NANA (Study 1) participants. Figures for age in Study 3 participants are averaged using mid-points in 5-year age ranges, so these contain a margin of error.

|  |  |  |
| --- | --- | --- |
|  | **Study 1** | **Study 3** |
| Number of participants with full datasets after data cleaning | 37 | 36 |
| Mean age (range) | 71.41 (65-89) | 71.02 (65-84) |
| Female/male | 21/16 | 24/12 |
| **Medical history:** | | |
| Participants with at least one medical condition | 29 (78%) | 25 (69%) |
| Participants taking at least one prescribed medication | 30 (73%) | 27 (75%) |
| **Smoking status:** |  | |
| Current smoker | 2 (5%) | 1 (3%) |
| Ex-smoker | 7 (19%) | 13 (36%) |
| Never smoked | 28 (76%) | 22 (61%) |
| **Factors affecting everyday living** | | |
| **Participants who reported:** |  | |
| Hearing difficulties | 7 (19%) | 7 (19%) |
| Difficulty reading newsprint | 2 (5%) | 2 (6%) |
| Difficulty preparing food | 0 (0%) | 1 (3%) |
| Difficulty shopping for food | 7 (19%) | 2 (6%) |
| **Technology usage** | | |
| **Participants who reported prior experience using:** |  | |
| A mobile phone | 35 (95%) | 36 (100%) |
| A digital camera | 29 (78%) | 34 (94%) |
| A microwave | 35 (95%) | 36 (100%) |
| Self-service supermarket check-outs | 20 (54%) | 35 (97%) |
| A computer | 34 (92%) | 36 (100%) |
| Participants reporting being regular internet users | 30 (81%) | 36 (100%) |

Figure 22 shows how mood scores were distributed across all participants over all days of the study. Higher scores are recorded for positive items (alert, happy, relaxed) while lower scores are typically recorded for negative items (hungry, sad, tired). The item ‘hungry’ has the greatest interquartile range, although the distribution of scores for the item ‘tired’ is also large relative to the other items.

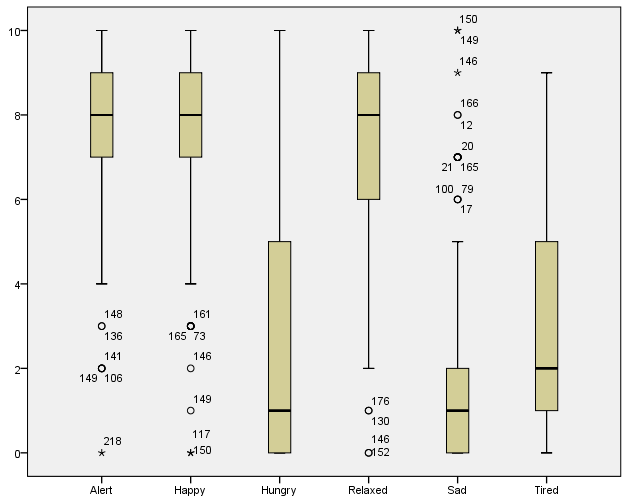


Figure 22. Boxplot showing distribution of mood scores across all days and all participants for Study 3

Comparing these distributions to those from Study 1 (see Figure 12, Chapter 4), we find a largely similar pattern, with the item ‘Sad’ having the smallest interquartile range and the lowest maximum score in both cases. However, the median scores for positive affect items (alert, happy, relaxed) are higher in the present study than in Study 1, and median scores for two of the negative affect items (hungry, tired) are lower in the present study. This suggests that participants in Study 3 recorded scores near the polarities of 0 and 10 more frequently than did those in Study 1.

Figure 23 shows the distribution of scores on the GDS at baseline and follow-up in Study 3. Similar to the results from Study 1 (see Figure 11, Chapter 4), the interquartile range of scores at follow-up was greater than that at baseline.

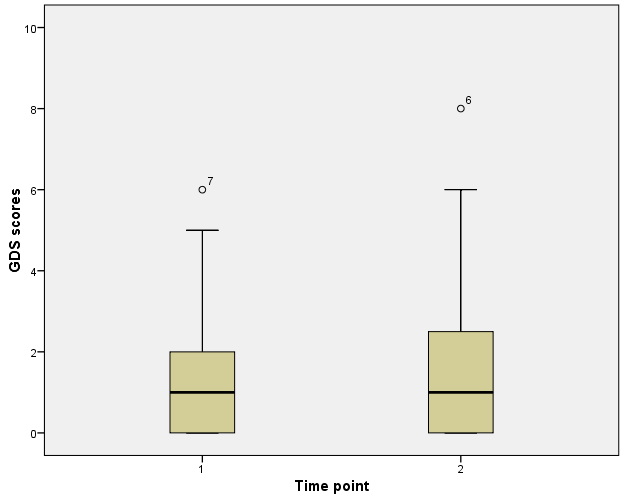


Figure 23. Boxplot showing spread of GDS scores for Study 3 participants at baseline and at follow-up.

Table 18 shows the numbers of individuals scoring above the cut-offs of three and five points on the GDS at baseline and follow-up. It also shows the number of individuals who crossed from non-depressed status to depressed status according to each of these cut-offs during the study. Fewer individuals increased their scores to above the five-point cut-off during the course of this study than did so during Study 1 (see Table 10, Chapter 4). Moreover, fewer participants scored over the cut-off of five points at follow-up in the present study than in Study 1. The known differences between the two samples may go some way to explaining this. A greater percentage of participants in Study 1 reported having a medical condition, and many medical conditions of later life are known to increase the risk of depression (Berkman et al, 1986).

Table 18. Frequencies of participants in Study 3 scoring above the cut-offs of three and five on the GDS (15-item) at baseline and follow-up

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **GDS cut-off used** | No. scoring over the cut-off at baseline | No. scoring over the cut-off at follow-up | No. scoring below cut-off at baseline, but above cut-off at follow-up | No. scoring below cut-off throughout the study | Total participants |
| **≥3** | 8 | 9 | 4 | 24 | 36 |
| **≥5** | 4 | 3 | 1 | 31 | 36 |

Figure 24 shows the kind of devices participants used to complete study tasks. The graph shows that the most commonly used devices were iPads and desktop computers. While these devices were preferred, other tablet computers, mobile phones and laptop computers were also used, with no fewer than four participants using each device listed.

Figure 24. Bar chart showing type of device used to complete daily mood reports. Participants were asked to select all devices that they had used.

### 6.3.2 Results from the ROC analysis

The following results are from the application of the predictive models generated in Study 1 to the mood and GDS data collected in Study 3. The model using a cut-off of five points used self-reported scores on measures of tiredness and sadness to predict future GDS depression status, while that using a cut-off of three used scores on measures of sadness and relaxation. Figure 25 shows the ROC curve from applying the model derived using a cut-off of five. The AUC is 0.69, and while this exceeds the chance level of prediction (0.50), the ROC curve for this result does not resemble the characteristic concave shape of a well-formed ROC curve. This can be attributed to the low number of participants scoring over five on the GDS at follow-up (3 participants).

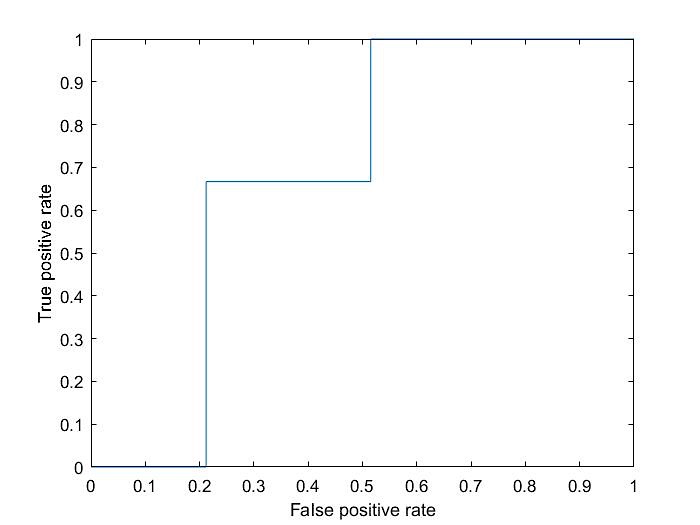


Figure 25. Graph showing ROC curve for the model at a cut-off of five. The AUC is 0.6869.

Since this ROC curve (Figure 25) is based on data from a sample, this represents an estimate of the ‘true’ ROC curve which would be generated if the dataset encompassed the whole population of (similar) older adults. Since there were few positive cases of depression, the ROC analysis provided an estimate of the curve, which was low in detail. Analysing a dataset which included a larger number of positive cases would increase the detail and therefore a more characteristic shape would be obtained.

Figure 26 shows the ROC curve for the application of the second predictive model, based on a cut-off of three points on the GDS. Since more participants scored three points or more than scored five or more at follow-up, the line takes on a slightly more characteristic shape. The AUC for this curve is 0.63, which is greater than the chance level of prediction at 0.5.

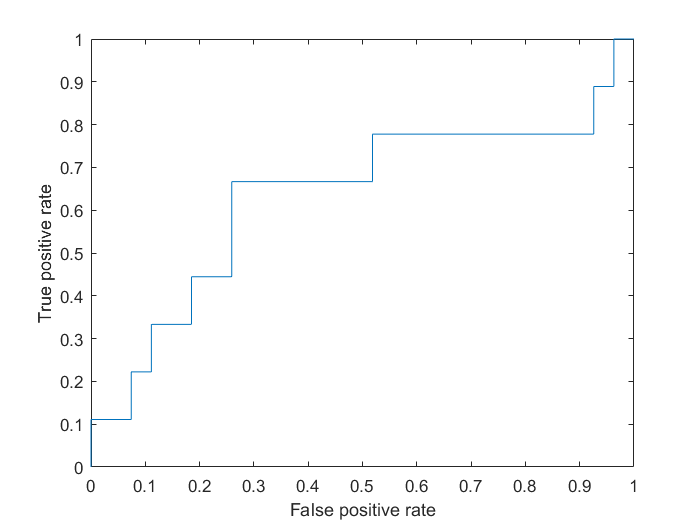


Figure 26. Graph showing ROC curve analysis for model using a cut-off of three points on the GDS. The AUC is 0.6337.

## 6.4 Discussion

This study explored the prospective validation of the models produced using LASSO in Study 1. The prospective validation of the models showed an above chance level of prediction for these models. This demonstrates the potential of using a machine learning approach with self-report measures of mood collected from older adults over multiple days.

Two models were tested prospectively, one using a cut-off of five points on the GDS, and one using a cut-off of three. The model with a cut-off of five points had an AUC of 0.69 when tested with the prospective data, although the ROC analysis produced a curve which was not of a characteristic shape. As a result, confidence in the validity of this curve is low. This occurred because very few participants scored above the cut-off of five at follow-up. When the model with a cut-off of three was tested, an AUC of 0.63 was obtained. In this case, a greater number of individuals scored above the cut-off, and this resulted in an ROC curve with a more characteristic shape.

The AUC can be used to compare the performance of one model against other predictive models to which ROC analysis has been applied. Chapter 3A reported a mapping review of studies which used machine learning models to predict future depression and anxiety. The majority of these studies were retrospective in design, while the study presented in this chapter is prospective. However, the mapping review did highlight one study which was prospective and used an ROC analysis, (Kessler et al, 2016). Kessler and colleagues’ study involved using the results of a one-time, self-reported depression screening test to predict the chronicity and severity of depression 10-12 years later. They reported achieving AUC values for their predictive models of chronicity and severity which ranged from 0.62 to 0.76. The AUCs from the study discussed in this chapter ranged from 0.63 to 0.69, and are thus comparable to the prospective study results presented in prior work.

Since none of the studies discussed in Chapter 3A addressed the problem of predicting depression in an older adult population, the present study has added to the research literature by demonstrating the potential of applying a predictive machine learning approach to data from older adults. The dynamic, self-report approach using digital technology over a number of days is also novel compared to existing approaches (see Chapter 3A). The variety in the devices used by participants in this study supports the finding in Study 2, that older adults have varying preferences in the types of technology they feel most comfortable using.

### 6.4.1 Limitations

Since this study explored the prospective validation of models to predict future depression, it was not possible to control for the number of individuals who would score above and below the cut-offs for depression at the end of the study. As it turned out, there were fewer participants scoring above the cut-off of five points on the GDS than was estimated in the sample size calculation. This may have resulted from the recruitment strategy – the majority of participants were recruited from the University of the Third Age (U3A), which is an organisation which facilitates the social engagement of older people. The fact that members are socially engaged through this organisation may mean that these individuals are less prone to depression, since depression is known to correlate with social isolation and loneliness (Alpass & Neville, 2010). The choice to recruit from this group may have meant that the sample had a lower incidence of depression than expected based on the rate of depression in the Study 1 sample. Future studies could use different recruitment methods in order to sample from older adults with a wider range of life experiences, including for example those who are more socially isolated. This may result in an increased number of cases relative to controls in the sample.

One limitation of this study was the use of self-report, which relies on participants completing the daily mood report and the GDS test in an honest manner. In Study 2, some participants described how reporting their mood on a computer may arouse fears concerning the consequences of reporting low mood, and that this fear may lead people to over-report positive mood and under-report negative mood (Chapter 5). This phenomenon may have affected the present study and that in Study 1 (Chapter 4), meaning mood reports were over-positive. An important dimension to this problem is whether participants who reported more positively on the daily mood report also reported more positively on the GDS. If participants were consistent in this regard, the validation of the algorithm itself would likely not be affected, since higher scores on the daily mood report would predict higher scores on the GDS. If participants were not consistent in this regard (i.e. consciously under-reporting negative feelings in the daily mood report, but completing the GDS more honestly, or vice versa), the models tested would likely not show good predictive ability.

Given that the models developed in Study 1 (Chapter 4) and validated here had above chance-level predictive ability, it is likely that either: participants did not over-report positive mood; or that participants were consistent in over-reporting positive mood in both the daily mood report and the follow-up GDS. Without interviewing the participants from both studies, it would be difficult to know for certain the extent to which this phenomenon took place. Future studies applying the same methodology could include a process evaluation in which participants were asked how honest they felt they were when reporting on their mood.

A future study could also collect data from a greater number of older adults, resulting in a larger dataset on which to train and test machine learning models. If areas under the ROC curve similar to those reported here were obtained from a larger training set, the approach could be applied in healthcare pathways to allow early intervention in the course of depression in older adults.

# Chapter 7 – Developing a model for the prediction of future anxiety in older adults

## 7.1 Introduction

Anxiety is a common mental health condition known to affect up to 15% of older adults (see Chapter 2). Anxiety disorders also commonly co-occur with depressive disorders (Nordhus, 2008). Anxiety differs from depression in that it does not cause an increased risk of all-cause mortality (Miloyan et al., 2016). However, similar to depression, anxiety is known to increase the risk of suicide (Chartrand et al., 2012), including among older adults (Kiosses et al., 2014). Aside from the increased risk of suicide, symptoms of anxiety are unpleasant, and are known to be associated with decreased quality of life in older adults (Bourland et al., 2000). Older adults with anxiety are also known to be greater users of healthcare than those without anxiety (de Beurs et al., 1999). Given the burden on older adults who experience anxiety, as well as the associated burden on healthcare services, early intervention in anxiety is likely to be beneficial.

In designing a tool to predict anxiety, it is important to consider how symptoms of the condition could be detected. In addition to excess worry, symptoms of anxiety include restlessness, irritability, sleep disturbance, being easily fatigued, as well as difficulty concentrating, muscle tension and elevated vigilance (American Psychological Association, 2013). Brown and colleagues have developed and validated six measures of mood and appetite with older adults (2016). Brown et al’s measures ask older adults to report, on a scale of 0 to 10, how much they feel each of six adjectives: happy, sad, tired, alert, relaxed and hungry.

These mood measures might reflect the symptoms of anxiety mentioned above. Restlessness and irritability might be detected by self-reported levels of relaxation, while sleep disturbance, difficulty concentrating, and becoming easily fatigued might affect the extent to which older adults report they feel tired or alert. The high comorbidity of anxiety with depression, which causes sadness and changes in appetite, among other symptoms (American Psychological Association, 2013), may mean that self-reports of happiness, sadness, and hunger are reflective of the onset of both depression and anxiety. The present study (Study 4) sought to explore whether the above-mentioned six mood and appetite measures could be used to generate a model predictive of future anxiety status when analysed using a machine learning approach, specifically the LASSO with logistic regression.

### 7.1.1 Research question

1. Can the LASSO with logistic regression be used to predict the onset of anxiety symptoms in older adults using measures of self-reported mood and appetite?

## 7.2 Method

### 7.2.1 Data collection

Data collection was conducted as described in Section 6.2.

### 7.2.2 Data analysis

Data cleaning was conducted as described in Section 6.2.6. Data analysis for the anxiety modelling was conducted in a similar way to the development of models described in Chapter 4: measures of happiness, sadness, tiredness, alertness, relaxation and hunger were averaged over one week’s mood reporting for each participant. Scores on the anxiety subscale of the HADS at follow-up were then coded as either positive (1) or negative (0) depending on whether participants scored above or below the threshold for an indication of anxiety on that test [Zigmond and Snaith recommend a cut-off of eight or above (1983)]. The mood scores were then used as input variables and anxiety status was used as the output variable.

Similar to the process used for the development of depression models in Study 1, the LASSO with logistic regression was applied using MATLAB. One hundred values of the selection parameter, lambda, were calculated in order to choose a level of coefficient shrinkage where deviance was minimized. A repeated, stratified cross-validation framework was used in the selection of lambda, using a five-fold approach as in Study 1. Two hundred Monte Carlo repetitions were used to avoid a favourable split. A coefficient plot was produced to visualise the results. The use of an ROC analysis was also envisaged to evaluate the predictive ability of the model.

## 7.3 Results

Figure 27 shows the coefficient plot for the model derived from the mood and HADS anxiety data using the LASSO technique. The blue dotted line indicates the point at which deviance is minimized. At this value of lambda, none of the mood variables have coefficients above or below zero, indicating that they do not contribute to the model at this point. As such, none were selected by the LASSO for inclusion in the final model. This indicates that these variables were not predictive of future anxiety status in this sample. Therefore, no ROC analysis was conducted in this study.

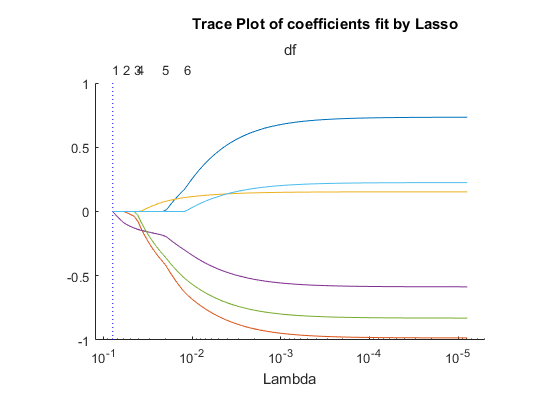


Figure 27. Coefficient plot for application of the LASSO to self-reported mood data and HADS anxiety scores

## 7.4 Discussion

This study explored the potential for the prediction of future anxiety in older adults by using the LASSO with logistic regression. The study aimed to derive a model which could successfully predict whether a user would score above the cut-off for anxiety on the HADS anxiety subscale. The LASSO did not select for any of the mood variables in the final model, suggesting that these variables were not predictive of later anxiety status. This contrasts with the Study 1 results where the LASSO was used to produce a model to predict depression status which included coefficients of multiple mood variables.

One possible reason that the chosen mood variables were not predictive is that the onset of anxiety may be more sudden than that of depression. Research has indicated that anxiety is triggered by certain life events or severe danger (Finlay-Jones & Brown, 2009; Angst & Vollrath, 1991). The fact that anxiety symptoms are triggered may indicate a more abrupt start to the experience of symptoms, such that presence or absence of these symptoms is more clear cut than in depression. This could reduce the chances that a self-report mood screen would be able to predict future anxiety based on self-reported mood.

Another possibility is that the words used to assess mood here are not as closely related to the experience of anxiety as they are to the experience of depression. Therefore using different words in a similar paradigm may generate a more predictive model for HADS anxiety scores. Anxiety scales typically ask users to report on feelings of nervousness, worry, restlessness and irritability (e.g. GAD7, Spitzer et al, 2006; HADS, Zigmond & Snaith, 1983; Pachana et al, 2007). As such, a study examining the use of these words presented in a similar paradigm to that used here may be more likely to provide a successful model for the prediction of future anxiety.

Chapter 3A of this thesis presented a systematic mapping review of studies applying machine learning to the prediction of future depression and anxiety. That review found only one study reporting application of machine learning to the prediction of anxiety exclusively (Devi & Kumar, 2016). In comparison, 12 studies were found that applied machine learning to the prediction of depression exclusively (see Chapter 3A for details). The comparative lack of studies applied to anxiety may reflect greater difficulty in the future prediction of this condition. Moreover, a positive publishing bias may mean that any studies that have been completed to predict anxiety but were not able to produce a successful model have not been published and therefore were not found in the searches. Alternatively, there may be a lack of interest in the early prediction of anxiety. However, the fact that anxiety is common (Bryant et al, 2008) and is correlated with increased risk of suicide (Chartrand et al, 2012) indicates that predicting this condition is still a worthwhile goal.

# Chapter 8 - Exploring the views of healthcare staff on use of technology for the prediction of future depression and anxiety in older adults

## 8.1 Introduction

This chapter describes the final empirical study. While Studies 1 and 3 explored the potential of applying a machine learning approach to the prediction of future depression and anxiety, and Study 2 explored older adults’ usability requirements and issues around motivation to engage with technology, the present study (Study 5) aimed to explore the potential implementation of predictive technology from the point of view of healthcare staff. The aim of this study was to explore the practicalities of implementing predictive technology with older adults within provision of community healthcare. The study took the form of a set of interviews with NHS staff who provide care to older adult patients at increased risk of anxiety or depression. The chapter begins by setting out the rationale, before stating the research questions, describing the methods, and presenting the results. These sections are followed by a discussion.

### 8.1.1 Rationale

In England, the cost of providing mental health support for those with long-term physical health conditions, is estimated at £8 billion to £13 billion per year (Naylor et al, 2012). Among older people, those with depression are found to have much higher costs of physical health care than those without depression (Unützer et al, 2009). Furthermore, as the population continues to grow, and people continue to live longer than ever before, healthcare settings are having to deal with increased demand and decreased capacity (Robertson et al., 2017). To address this, the NHS England Five Year Forward View strategy document discusses the need for targeted prevention of avoidable conditions, as well as empowerment of patients through the use of digital technology (Mental Health Taskforce, 2016). Early intervention in episodes of depression is known to improve outcomes (Sirey et al., 2005; Reynolds et al., 2014), and, as such, use of low-cost technology managed by the individual in their own home to detect early signs of depression and anxiety has the potential to help alleviate the distress caused by these conditions as well as the associated economic burden.

Involving healthcare staff from a range of levels in the development and implementation of new technologies is known to be important for successful adoption and sustained use in practice (Ackerman et al, 2012; Kyratsis et al 2012). As such, this study seeks to explore staff's views of the practicalities of using a touchscreen application to detect early signs of mental health problems in older adults.

Community healthcare staff were chosen as the population of interest in this study because their roles involve working with older people with long term conditions known to increase the risk of depression and anxiety. This choice was informed by conversations with healthcare staff at two public engagement events, which were organised by the researcher and three colleagues in 2015 and 2016. These events involved bringing together older adults, NHS staff, technology companies, charities and academics to discuss the use of technology to support mental health in later life. These events provided opportunities for the researcher to discuss his research with NHS staff and others prior to conducting the present study.

### 8.1.2 Research Questions

The research questions for this study were as follows:

1. What pathways exist in community healthcare services in Sheffield for the detection and management of depression and anxiety in older adults?
2. What are the views of community healthcare staff on the use of digital technology with older adults for the prediction of depression and anxiety within these pathways?

### 8.1.3 Approach

The study took a qualitative approach, since the aim was to explore views and perceptions among healthcare staff. Marshall and Rossman (2006) describe qualitative methods to be useful in “research that delves in depth into complexities and process” (p53), and in “research on little-known phenomena or innovative systems” (p53). This approach was deemed appropriate for the aims of the present study, since community healthcare pathways are complex, and involve multiple processes. Meanwhile, the novel application of mood-reporting for prediction of depression and anxiety can be considered an innovative system.

## 8.2 Methods

### 8.2.1 Setting

This study took place within the community health section of one healthcare trust which works across Sheffield. Interviews were conducted in August 2017. Six months prior to the start of the study, the trust had undergone an organisational change around its community services. Sheffield Teaching Hospitals was part of the second wave of pioneer trusts where community care services were integrated together, under a government scheme titled the ‘National Collaboration for Integrated Care and Support’. This had led to the merging of three, previously separate, community teams (active recovery rehabilitation, falls prevention, physiotherapy/occupational therapy) into one single team, charged with the rehabilitation and healthcare needs of housebound patients in the community. Twenty-five pioneer sites across the country have undertaken the integration of community care services, including Sheffield. As such, the situation in Sheffield experienced by the participants in this study is likely to reflect current care in the other 24 pioneer sites.

Care for depression and anxiety in the trust from which participants were recruited follows the stepped care model in the guidelines set out by the National Institute of Health and Care Excellence (NICE) (NICE, 2011; 2016). The stepped care model has five steps, each recommending a different course of action depending on the severity of a patient’s condition. Steps one and two are designed to address low to moderate depression and anxiety, while steps three to five deal with moderate to severe cases.

### 8.2.2 Participants

Participants for this study were staff members of the Sheffield Teaching Hospitals trust who had worked with older adults known to be at increased risk of depression or anxiety in the community for at least one year. Healthcare staff were asked to self-certify whether or not they met the inclusion criteria. Participants were purposively sampled from Sheffield Teaching Hospitals trust to select staff members from multiple levels of seniority (or ‘bands’ as they are termed within the NHS) and multiple roles. Participants were sought from among both management and frontline staff, and from staff with different specialisms (e.g. occupational therapy, community nursing, management). The sampling of staff from a variety of roles, specialisms and levels ensured that the feasibility of implementation of the technology was practically addressed from multiple angles, which is known to be important in implementation research (Mohr et al., 2017a; Kyratsis et al., 2012; Ackerman et al., 2012).

The researcher worked with a principal investigator within the NHS to identify and recruit suitable healthcare staff to interview. Twelve members of staff were invited to take part. These included general practitioners, community nurses, physiotherapists, occupational therapists and service managers.

### 8.2.3 Ethics

This study had two main ethical issues. Firstly, the study sought participants’ views and opinions, and in giving their opinion, it was possible that participants may have criticised the way services were organised or managed. Thus it was deemed important that participants were assured of the anonymity of their responses. Anonymity was achieved by removing personally identifying information from data at the stage of transcription, and ensuring that participants were not identifiable in written reports generated from the study data.

Secondly, the study involved discussion of mental health conditions, and this may have caused upset for any participant with a close connection to someone with an experience of such conditions. As such, the discussion of mental health conditions within interviews was conducted with respect and sensitivity.

The study received a favourable review from the ScHARR Research Ethics Committee at the University of Sheffield, as an amendment to ethics application number 003140 (see Appendix 5). In addition, research governance approval was granted by the Health Research Authority (project ID 223998 and REC reference 17/HRA/3130; see Appendix 7).

### 8.2.4 Procedure

After reading the information sheet, completing a consent form and completing a short demographics questionnaire, participants were interviewed using a semi-structured, in-depth interview approach. A semi-structured interview asks several main, standard questions to all participants, and follows these questions with on-the-spot questions to further explore comments made by participants (Jamshed, 2014).

The interview schedule used in this study had five main questions. These included questions about the participants’ job role and experience using technology or telemedicine with older adults. The schedule also included a description of how self-reports could be used with digital technology in the prediction of future depression and anxiety with older adults. Participants were asked about their immediate reactions to this idea, as well as about barriers they perceived to its implementation. The full interview schedule is presented in Appendix 8.

Interviews each lasted 20 to 50 minutes and were recorded using an encrypted voice recorder. Recordings were then transcribed verbatim on a computer for analysis, at which point, all identifiable data were removed. Interviews were conducted at each staff member's place of work.

### 8.2.5 Data analysis

Thematic analysis (Braun & Clarke, 2006) was used to analyse the data, and this involved several steps, or phases: i) familiarisation with the data, ii) generation of initial codes, iii) finding themes, iv) reviewing themes, v) defining and naming the themes, and finally, vi) producing a report (Braun & Clarke, 2006).

Here, familiarisation was completed by transcribing the recordings, and re-reading these a number of times. Data were coded in the computer software package Nvivo (QSR International), which enables coding and organisation of codes using a hierarchy of ‘nodes’.

The process of generating codes involved identifying a unit of meaning within a transcript, and either: assigning it to a new code in Nvivo, which was then given a short description; or assigning it to an existing code, the description for which adequately described what was identified in the new section of text. Coding was conducted by the researcher alone.

To find themes, the descriptions of all codes were first read through. Similarities between multiple codes were then identified and similar codes were amalgamated into thematic units. These either became full themes in their own right, or were treated as subthemes and matched with similar subthemes into a larger overall theme. To give an example of the process of finding themes, the codes ‘Not enough mental health care staff’, ‘Staff motivation’, and ‘Need staff to monitor incoming data’ were all considered to relate to the underlying issue of management of staff, so these were joined under a thematic unit called ‘Considerations around staffing’. Other thematic units included ‘Cost’ and ‘Challenges for getting timely and accurate data’, and these were joined with the aforementioned thematic unit to produce an overall theme on ‘Practical considerations around implementation’, with each of the thematic units as subthemes.

The thematic structure was developed iteratively, and involved reorganising the themes and codes multiple times to obtain a structure that adequately represented the data. This process was conducted by the researcher alone.

The resulting set of themes, subthemes and codes were then reviewed and discussed by the researcher and one member of his supervisory team. As a result of these discussions, the themes were slightly reorganised to reflect the breakdown of topics among the research questions.

## 8.3 Results

### 8.3.1 Participant information

Eight members of staff showed an interest in taking part, and seven were finally interviewed because one became unwell and was not able to rearrange the interview. A breakdown of the job role of each participant is provided in Table 19. Five of the interviews were conducted on a one-to-one basis, while one of the interviews was conducted with two participants at once, owing to their limited availability. The sample included two participants specialising in physical health care, one participant specialising in mental health care, two participants who had been dual-trained to carry out both physical and mental health care, and two service managers, both of whom had prior experience of working in the physical health care of older adults. All participants were female.

Table 19. Job roles and experience of participants in Study 5

|  |  |  |
| --- | --- | --- |
| Participant identifier | Job title | Time in current role |
| P1 | Community mental health nurse | 0-2 years |
| P2 | Community staff nurse | 0-2 years |
| P3 | Registered general nurse/Dual-trained practitioner/Psychological wellbeing practitioner | 2-5 years |
| P4 | Community nurse/Dual-trained nurse | Over 10 years |
| P5 | Operation manager for ICT-Therapy | 0-2 years |
| P6 | Integrated pathway manager | 2-5 years |
| P7 | Occupational therapist | 5-10 years |

### 8.3.2 Thematic Analysis

The thematic analysis generated six main themes to address the research question for this study. These covered: suitability of the integrated care team as a host for predictive technology; current pathways of referrals and information exchange; precedents in the use of technology with older adults in community health care; benefits of implementing predictive technology; challenges for implementing this technology with older adults; and challenges around staff engagement. Findings from each theme, including example quotes, are presented below.

**Theme 1: Suitability of the Integrated Care Team as a host for predictive technology**

The first theme included quotes that characterised the work of the Integrated Care Team. This contributed to the research by establishing an understanding of the type of patients seen by the service, and of how depression and anxiety were currently detected and managed by the service. Participants explained that the ICT coordinates the care of patients with a variety of healthcare needs, many of which predispose them to a higher risk of common mental health conditions.

P5: The patient cohort that we have, are, anybody over the age of 16 with a Sheffield GP, and we can have a whole wide range of patients, orthopaedic, respiratory, muscular-skeletal, fractures, frailty, fallers, the caboot so, some of the patients will have obviously anxiety and mental health issues.

Participants expressed that, while the ICT service was technically available to anyone over the age of 16, they mainly worked with patients over the age of 70.

P6: the age range is from 16, 18, but the majority of our patient group would fall in their 70s, 80s, 90s.

While most of the roles carried out by staff in the ICT were physical health roles, participants also recognised the importance of mental wellbeing in overall health.

P5: obviously mental health has a big effect, as a physio, on people, living, and it’s key to keeping people well both mentally and physically, so the more we can do to keep both ends sort of up, the better.

Participants reported that occupational therapists used two screening tools to detect mental health difficulties in new patients, these being the six-item cognitive impairment test (6CIT) and the four-item geriatric depression scale (4GDS). Further tests were used if these two short scales indicated a problem with mental health.

P7: the 6CIT and the 4GDS tend to be done at the initial assessment, and if they trigger on those, we then do, I do a PHQ9 and a GAD7, for the anxiety and depression.

However, participants admitted that recent organisational changes and competing priorities meant that these procedures were sometimes neglected, and increased patient numbers in the newly formed service have increased pressure on staff-patient contact time, reducing what is possible during home visits.

P7: up until April this year, it was routine that all the new initial assessments, we did that, to rule out, or pick up on any sort of mental health problems, [..] May or April this year, a lot of the services merged, things are a little bit different, so it’s not sort of stated that we have to do it now, but it was before [..]so we probably only tend to do it if it’s picked up on.

R: If it’s picked up on? (P7: yeah) So if, if there’s, let’s say a physical indication? (P7: Yeah)

P6: it was initially in their [occupational therapists’] remit to always do a mental health screen for every patient, but as we’ve merged the services and increased the number of patients, [..] it becomes a bit unwieldy to be doing it for every single patient

**Theme 2: Current pathways of referrals and information exchange**

The second theme included quotes demonstrating how depression and anxiety in older adults were currently detected and referred between staff roles. Participants explained that in Sheffield, patients at steps one and two of the stepped care model are referred to occupational therapists, who provide low level interventional support.

P1: This team works in a stepped model, so we kind of have steps 1, 2, we would expect, kind of OTs or other members of staff to be able to manage that, so that’s kind of low mood, but fairly new, and that they could do some interventions to help that not carry on or not to get to the point of depression that’s more severe

In the quote below, a service manager discusses procedures in place if a community nurse identifies a need for mental health support in a community patient.

P6: If they identify that there’s a need, they would refer it to the OTs [occupational therapists], who would start that initial, low level interventions, and if it was more than the OT could deal with, it gets referred to the mental health nurse.

For steps three to five in the model, patients are referred to community mental health nurses who provide additional support and signposting. An occupational therapist described how she and others in her team assessed patients to determine whether they required the attention of a mental health nurse.

P7: We would do, depending on what it was we would do the GAD7, the PHQ9 or maybe a mini mental state exam, and then we would look at signposting on from that

A community mental health nurse working within the ICT unit described the kind of support they provide.

P1: we do do some kind of treatment with people, and again that can be, maybe people who are a bit more risky, or who are a bit higher up in the stepped care model, again, looking at goal-setting, problem solving, motivation, and trying to start that off, and then deciding, monitoring over time how that works, how they respond to maybe, self-help literature, or the suggestions and then that can be then passed back to the GP to be carried on, or be referred on to another service to pick that up after we’ve finished.

**Theme 3: Precedents in the use of technology with older adults in community healthcare**

This theme included participants’ comments on existing uses of technology for the remote monitoring of older adults’ health status. Participants gave examples of two technology-based systems already in place that some older adult patients had used to report on (physical health) symptoms remotely. The first of these was referred to as ‘Telehealth’. Although the term ‘Telehealth’ is used within the research literature to designate a category of technologies (see Chapter 2), participants in the present study used the term ‘Telehealth’ to refer to a particular system that they had used. Therefore the word ‘Telehealth’ and the phrase ‘the Telehealth system’ are used to refer to this particular system within this chapter (the true model name of this device is unknown to the researcher). The Telehealth system was used by older adult patients to report measures including blood oxygen saturation, taken using a pulsoximeter, and blood pressure, taken using a blood pressure cuff.

R: Mm, and that included, for example a blood pressure machine, or something?

P6: Yes, so you would be able to attach your blood pressure cuff to it, and, I don’t know whether the pulsoximeter was attached to it or separate, but it was all connected in some way, yeah.

Patients reported their readings onto the Telehealth system, and these readings were transmitted as data to the community care staff via a telephone line. A system was in place where nurses would be alerted if a patient’s data crossed particular thresholds. Participants described how some members of staff were tasked with monitoring the incoming data and dealing with alerts where necessary.

P6: that information goes to the single point of access, and they look on the, I think it’s a website, to check the readings, and then they alert the nurses, if the readings are particularly low. [..]

Researcher: And who monitors the single point of access?

P6: So there are nurses and admin in our single point of access, from something like 8am to 10pm, or something like that, 7 days a week.

Researcher: So someone is seeing a read out there, every, every day, what’s coming in.

P5: Yeah, and it would alert if it was missed, or if it was outside the parameters that had been set.

These quotes describe pathways that are already in place to manage data coming from patients. Participants described too how training was provided to patients who were invited to use the system.

Researcher: Mm, and so it relied on the patient using that by themselves?

P6: Yeah, so we had somebody that would go out and train them how to use it.

Where participants struggled with the system, contingency plans were in place to ensure care was not affected.

Researcher: And when, so when someone didn’t cope with the technology, were you, was the team ready then to step back in and say ok we can take over this?

P2: Yes. It would just be taken out and say, you know you would say, oh that’s fine.

Participants spoke of how the Telehealth system benefited patients in unexpected ways.

P5: it was surprising how a lot of the older adults quite enjoy it. (R: right, ok) There are some that struggle, maybe with dexterity, but there was a lot that took us by surprise, who enjoyed being monitored

Some participants suggested that predictive technology could be implemented within the same Telehealth system.

P2: So that could be used in, with Telehealth, couldn’t it? So they could do their inputting, or, however they, answer the questions could be through Telehealth, as well, if they’re not seeing us every day.

The second example of technology currently in use with older adults in the ICT unit was a system called ‘Florence’, abbreviated to ‘Flo’. One participant described the uses and benefits of the Flo system.

P6: Some of it has been about monitoring people’s oxygen saturation levels, some of it might be blood pressure, or it might be prompting them to do an exercise programme.

P6: We’re trying to increase the use of Flo within our community services, the idea being that you know, the patient can feed back information which can be monitored, and it’s preventing a nurse or a therapist having to visit as frequently [..] what I think it tends to do is um, provides a level of monitoring that you can react to.

In contrast to Telehealth, Florence was described as operated via text message, and offering a level of automation.

P6: the patient can respond, they might, there might be more questions so that might then prompt another text, so there can be a bit of interaction, but it is, automated, it’s automated responses, and that kind of thing.

R: Ok, so there’s nobody sat there in a desk, (P6: No) sending texts to different people?

P6: No

While all participants were aware of Florence, few had experience using it, so fewer comments were made about this form of technology than about the Telehealth system.

Researcher: have you been involved with Florence?

P7: I’ve heard about Florence, but I’ve not had anything to do with it.

**Theme 4: Benefits of implementing predictive technology perceived by staff**

The fourth theme covered the benefits that participants perceived a predictive system for depression and anxiety might have. For instance, participants described how they believed predicting future conditions could reduce the burden of treating people at later stages of illness.

P1: I think the idea of picking it up earlier though, probably means that it would be easier to manage

P4: I think it’s ideal because you catch these people before they tumble down into a downward spiral

P3: if we can pick that up in the very early days, we can really, it might only be a bit of psychoeducation [..] I think there’s massive health gains and savings to be made, if we can start routinely screening and signposting people to the right help at the right time.

Participants also suggested that using self-report scales on digital technologies could be empowering for patients. For instance, one participant described how patients could benefit from having control over reporting their mental state.

P4: They get dignity, they get control, by identifying where they are at, what’s happening to them, giving them the means to move themselves forward, which they wouldn’t have had in the past.

This participant also discussed how patients may feel keener to engage through technology since they would not be taking up clinician’s time:

P4: it gives them an added reassurance, they’re not bothering clinical people

Another participant commented on how technology could facilitate mood reports where patients may have struggled with pen and paper mood diaries:

P1: Yeah I think we have done kind of mood diaries with older people as well. I think that the very old and more physically unwell people might struggle to do that, and often it’s a case of being able to write things down, and track it, and that way can be very difficult, so if people have an easier way of doing that, that might be useful.

Using technologies familiar to older adults was seen as a way to overcome some of the barriers associated with implementing technological interventions, for example using text messaging.

P4: text is a really good way, by phone, and they feel quite happy with it, that they’re not struggling to listen, and they can sit and read, and if they can’t read it, they can go to the neighbour and go ‘what does this mean then’.

Thus, participants recognised several benefits that a digital self-report prediction system could bring to community care.

**Theme 5: Challenges to implementing this technology with older adults**

Participants perceived several potential barriers to the successful implementation of prediction technology with older people. Many of these were concerns regarding the age of the target population. A particular area of concern was cost. In participants’ experience, older adults were often concerned about the cost of running technologies provided to them.

P3: And the cost of the patients using it, whether it is electricity or, or they’ve got, if they’re using a mobile, have they got to top up

P4: they just worried about breaking it, and how expensive it is, those sort of things, and how much electricity it’s using, is a big thing. How much it’s costing.

Participants discussed the cost of digital technologies where participants do not currently own them, and described how this could be a barrier to adoption if the cost of the devices fell on the health service. This was particularly of concern in managing the care of older people, since they were less likely than younger adults to own digital devices.

P6: With the sort of, screen thing that you’re talking about, there would be things like the cost to the organisation [..] who pays for it?

P5: technology as such, we have very limited… Predominantly because of cost.

P3: Is there a cost for us to buy technology for them to use, whatever kind of technology that is?

Participants mentioned that finding money to pay for these would depend on funding streams within the organisation.

In addition to cost, participants discussed a number of other challenges around the use of technology by older patients in their own homes. This included that any supplied technology may be lost, broken or stolen.

P7: It’s how safe would that be in someone’s house?

Researcher: Oh right ok, in terms, safe as in, to not get broken?

P7: To not get broken, or to not get stolen, or just get mislaid in general.

Another participant raised the issue of infection control, where devices are taken between patients’ houses.

P6: there’s infection control, because we’re taking it from patient to patient.

Participants also commented that the requirement for either mobile reception or wifi should be considered.

P6: how does it report back? So does it, is it through that mobile phone technology, or does it rely on wifi, all of these kinds of things?

While the supply of technologies involved difficulty, challenges were also perceived in getting older adults to successfully engage with technologies once supplied. Getting older patients to engage with technology at all was seen as a challenge by an occupational therapist.

P7: I think in theory it’s a good idea, I think trying to get people to engage with that, may not be so easy, just because of the technology.

The interviews explored the reasons participants believed using technology with older adults may be a challenge. One participant reflected on a piece of technology she had used with older adult patients in the past, a machine for monitoring blood oxygen levels. She suggested that some patients did not engage well with it because they did not understand or feel comfortable with the technology.

P4: Some people, I found didn’t like it, just didn’t get on with it, and that just, everybody’s different, everybody’s an individual with what they can cope with and what they can understand, and I think it’s the approach really, with technology.

Some participants expressed how, in their experience, it was the set-up of new technologies that posed the greatest challenge.

P1: they might say, I like to, you know, watch videos on this, but my family helped set it up. It’s the practicalities of setting it up and once they know how to do it, it’s a bit easier.

Memory was another reason cited by participants for older adults’ disengagement from digital technologies. Prompting was seen as essential for users to remember to complete required tasks.

P2: Memory is another one [potential barrier], remembering to do it, if they’re not prompted, how are they gonna remember to do it?

P5: I would have concerns about people doing it every day unless they were prompted, because from trying to get people to do exercises every day, it’s a real battle, so I think they would need some kind of prompt

P7: you could almost do with something on it, that beeps or like an alarm, to remind people, because that’s another thing is just memory, will people actually remember to do it?

Some participants suggested possible procedures for prompting, although there were concerns around staffing to implement these procedures.

P5: whether it’s a telephone call to say ‘have you done it?’, or whether it’s somebody actually saying, come on, have you, let’s do it today. I don’t know.

P7: I mean, I suppose, a telephone call, could be done, at a set time every day, but it would be, who’s gonna take that, who’s gonna do that call, and it’s not possible to guarantee that on a set day you’re gonna be free to do that, so that would probably be quite difficult to set up.

Use of prompting was not seen as a perfect solution. One participant had experience using an app for mental health support with younger patients which featured alerts. In her experience, these prompts could be annoying or upsetting for patients.

P3: a couple sort of like said, ‘well actually, it kept reminding me, at sort of like inconvenient times’, things like that, or one person said it just made them feel, it put them in a bad mood, because they didn’t want to be reminded of how they were and it sort of like reminded them.

Thus, while a system of alerts or prompts could help to promote timely use of self-report technologies, consideration should be made for how these prompts are conducted, taking into account the impact the prompts themselves, as well as the information requested, may have on users.

**Theme 6: Challenges around staff engagement with predictive technology**

The sixth theme included views on staff engagement and how this might impact on the implementation of predictive technology. Participants had varying opinions on whether staff would have sufficient time to set up, monitor, assess and respond to new technology for prediction. Participants who worked in frontline staff roles believed that there would be sufficient time to set up and introduce technology to older adult patients.

Researcher: Do you think that you’d have time to train someone to use that sort of, say a very simple app?

P7: I don’t see why not. It might be that the initial assessment would be extended, but if that was part of it, then that’s fine, do you know what I mean, it would just be included and there’s no reason why it couldn’t be.

Researcher: Would there be similar sort of periods of time where you could see that fitting in?

P2: Yeah, I don’t see why not.

Researcher: Could you give any examples or perhaps describe a bit what that might look like?

P2: Well I could be talking to a patient about how they feel, which is what I tend to do all the time anyway. And if they’ve got something in front of them they can answer in it in that way rather than just talking to me.

One service manager was concerned about the amount of time it would take to train the patients, and how many would need to be trained, but believed it would be possible.

Researcher: Do you think there’s capacity within the teams to provide that training to the patients? Do you think they’d have time to do that?

P6: So it would depend upon the amount of training that was needed. There are, so within the teams we have different grades of staff. And sometimes if the clinician doesn’t necessarily have the time, they could arrange perhaps for the support worker to go on and spend some time with the patient, so there is the potential there, but it would come down to the numbers

However, participants expressed concerns regarding staffing for monitoring the data generated by the system and following up with full mental health screens if the system predicted the incidence of depression or anxiety.

P4: The community nurses are overworked, who’s going to do the full screening [..] What do you do with the information? Who’s gonna monitor, control it, who’s gonna read it?

Participants also expressed concern about the requirement for staff to follow up with patients who were predicted to score highly for depression or anxiety. Several participants believed staff would avoid conducting preventative screening practices if there was the possibility it could lead to extra work.

P5: With the pressures that we’re under, predominantly to get people out of hospital, to get them home and to keep them safe, and then the other end prevention, it would be another piece of work really, and it’s not so much answering it, it’s then doing something with that knowledge isn’t it

P3: Once you’ve screened for something, you’ve got a responsibility to follow that through and do something about it, otherwise it’s neglectful. And I think that’s actually…

P4: So if you don’t ask, you don’t have to follow it up. And that depends on the professional really, as well.

Some expressed this as a need for a change of culture, whereby staff needed to understand the benefits of preventative systems.

P6: There’s a whole cultural change that we’re going to have to go through, in terms of getting people on board, and seeing it as part of their role, and you know, them thinking that they’ve got the time to do it, and understanding that actually by dedicating some time here, you might save some time down here

## 8.4 Discussion

### 8.4.1 Principal results

This study used semi-structured interviews to explore the views of community healthcare staff on the use of digital technology for predicting common mental health conditions in older adults. Seven participants, from a range of roles and levels of seniority in a community healthcare organisation were interviewed. Six themes were developed from the transcripts, using Braun and Clarke’s Thematic Analysis (2006). The first theme explored the suitability of the proposed system to community healthcare, while the second theme explored pre-existing pathways for the management of depression and anxiety in the integrated care team. The findings from these two themes highlight how the service examined in this study is already equipped to deal with early signs of common mental health conditions, and already takes a pro-active approach to the prevention of future crises through the provision of low level support from occupational therapists. The existence of these pathways means the introduction of a predictive technology system might fit into existing pathways within the ICT unit. For example, predictive alerts could be forwarded to occupational therapists who could begin treatments advised at the lower ends of the NICE stepped care model, as per their current remit.

Theme 3 examined how some technologies for older adults to self-report health data are already in use in community healthcare. Although these technologies focus on physical rather than mental health symptoms, the successful implementation of Telehealth and Florence could provide a template for the implementation of predictive technology for depression and anxiety. Telehealth (and more recently, Florence) allow patients to take certain measures of physical health (e.g. blood oxygen, blood pressure) and submit these to healthcare teams digitally, via a phone line, or by text message. Staff explained that services are organised around these systems in such a way that participants’ submitted measures are monitored by a member of staff at the ‘Single Point of Access’. Any deviation in these measures from pre-set parameters raises an alert, and the readings are referred to an appropriate member of the care team. A similar model could be employed for the self-report of mood data. This data could be submitted to the Single Point of Access, where an alert could be raised if a patient’s mood scores are indicative of future depression or anxiety, according to the model developed in Chapter 4. Further research would be required to assess this proposed pathway in practice.

Theme 4 examined participants’ views of the potential benefits predictive technology could bring. Benefits were perceived for the health service, including that early interventions could save both time and cost in the future, and also for patients, who participants believed would gain a sense of empowerment through using such a tool. A cost-effectiveness study would be required to understand the true value to the organisation that the proposed solution could bring.

Themes 5 and 6 examined the challenges associated with implementing predictive technology. Challenges mentioned by participants included the cost and supply of digital technologies, while patient engagement with the technology was also raised. Theme 6 explored issues in staff attitudes and engagement with the proposed approach. Participants mentioned lack of time and the perception of an increased workload as being potentially problematic for staff. Although findings in Themes 1 and 2 showed there to be potential for the implementation of predictive technology within existing pathways, the findings in Themes 5 and 6 indicate that there would be a number of challenges to overcome before the approach could be implemented successfully.

### 8.4.2 Comparison with prior work

This study complements previous work on implementation and adoption of healthcare technologies, as well as work in other areas. The finding that participants saw self-report healthcare technologies as empowering for patients (Theme 2) reflects work by Mitzner et al (2010) which showed that older people value the convenience of technology for supporting their healthcare needs. While Mitzner et al’s work only explored the views of older adults themselves, the work here highlights the fact that health professionals too are aware of this.

Cost of purchasing and running technologies, discussed in Theme 5, is a common theme in literature around older adults’ views of technology in general. Greenhalgh et al (2010) and Peek et al (2015) both discuss older adults’ concern about this aspect of technology usage. Ofcom has reported (2015) how in the UK, uptake of smartphones and tablet computers is growing rapidly among older adults. As such, the expense of providing digital technologies to older adults may be reduced within a small number of years, as these become more ubiquitous. Costs associated with implementing new technologies have also been raised as an issue of great importance to healthcare providers in prior work (Kyratsis et al, 2012). An economic analysis would need to be carried out to estimate the cost and savings to health organisations of implementing predictive technology for depression and anxiety. This would allow recommendations to be made as to whether it is in the interest of a health organisation to purchase and provide digital technologies for those older adults who do not yet own a digital device, to self-report mood.

Theme 6 discusses how staff may be reluctant to instigate new detection methods if doing so may require to them to take further action later on. Similarly, Ackerman et al (2012) describe how nurses were reluctant to make use of a self-diagnostic kiosk for urinary tract infections, since it was found to create more work to manage patient care. Ackerman and colleagues (2012) recommend the use of ethnographic research methods during the implementation of new technologies to understand from a sociotechnical perspective the quandaries and unexpected occurrences in such implementation work.

Continuing to apply research methods throughout the implementation of a new technology in healthcare settings may offer a way to ensure implementation is sustainable. Mohr et al’s Accelerated Creation-to-Sustainment model (ACTS; 2017a) advocates participatory design at the development stage of new technologies. Their model also advocates “iterative processes of evaluation and design to occur during the trial phase [..] as well as semi-structured user feedback interviews”. Mohr and colleagues argue that this approach can help to achieve successful implementation of new technologies (2017a). Any future studies of the implementation of the predictive technology approach described in this thesis could include an element of ethnography, as advocated by Ackerman et al (2012). Future studies could be used to evaluate and redesign the technology, or the pathways used around the technology, in an iterative process, as described by Mohr et al (2017a). This may help to sustain future use of the approach.

### 8.4.3 Reflexivity

Braun & Clarke (2006) describe the need for researchers to be reflexive when conducting thematic analysis. This section therefore discusses how the researcher’s background and interests may have influenced the research completed.

The researcher conducted this study as part of a thesis on the use of technology to predict depression and anxiety in older adults. The first and third studies presented in this thesis (Chapters 4 and 6) each relied on the researcher to make a positive effort to make the technology work, for each stage to be considered a success. As such, the researcher may have been influenced, consciously or unconsciously, by this ‘need to succeed’ in the present study. This could have resulted in unfairly biasing the analysis towards positive views of the technology discussed.

The stages at which bias could have been introduced were: i) sampling; ii) development of interview questions (both standardised questions and follow up questions); iii) coding the data; and iv) presenting the data in the written analysis. However, the researcher remained aware of the potential for bias throughout, and mitigated the risk in the following ways. i) The initial sampling of healthcare staff in the ICT unit was conducted by an independent member of staff in the healthcare trust who took the role of principal investigator. This individual did not have any vested interest in the outcome of the study, and so was less likely to have been biased in her choice of invited participants. ii) The standardised interview questions were reviewed by the researcher’s academic supervisor as well as by the Health Research Authority. These individuals were less partial to the outcome of the research and so could provide a critical eye. During the interviews, the researcher maintained a conscious awareness of the possibility of leading questions, and strived to ensure questions were fair. iii) The codes and themes generated in the analysis were discussed with a member of the supervisory team who had expertise in conducting thematic analysis. iv) The write-up of the study was reviewed by other researchers who were less partial to the outcome of the study.

### 8.4.4 Strengths and limitations

The main strength of this study is that the seven participants interviewed were drawn from a broad range of staff roles and levels. All participants matched the inclusion criterion of having one or more years’ experience working with older adults at increased risk of mental health conditions, meaning participants had a practical knowledge about the topics discussed. The varied perspectives that members of staff brought from each role served to make the data richer, and strengthened the findings where commonalities between their perspectives existed. For example, the fact that both a service manager and a community nurse described cultural change as necessary to introduce preventative practices, suggests that this knowledge exists across levels of seniority.

There are also some limitations to the study. The sample was taken from only one healthcare trust in one English city, which means some views may have been only relevant to the context of that particular trust. However, given that 24 other regions have, like Sheffield, integrated their community healthcare services, it is likely that the findings here are relatable to the situation in these regions as well. Furthermore, the findings presented here are supported by prior research in other settings, and it is therefore likely that this work can provide some indication of where challenges may lie more generally. For example, challenges found here included cost to the user, and a shortage of healthcare staff time for the administration of technology and newly created data. Given older adults’ known concerns about the costs associated with technology (Mitzner et al, 2010) and given also the ‘productivity challenge’ currently faced by the NHS (Appleby et al., 2014), it is likely that these challenges will be common to many healthcare services in the UK.

The sampling strategy for this study involved purposive sampling through the principal investigator, who was an employee of the healthcare trust where the study was conducted. The principal investigator had been briefed on the project being conducted, and it is possible that the choices she made of who to invite to take part in the study were influenced by her perceptions of who would give preferable answers. Further sampling bias may have occurred because participants were informed about the subject of the interview before consenting to take part. It may be that only those who endorsed a positive view of the deployment of technology in healthcare therefore expressed an interest in taking part. This may have meant that responses to interview questions in the final dataset were more positive about implementation of predictive technology than responses may have been, had other members of staff taken part.

Another limitation is that the time available for interviews to be carried out, along with limited healthcare staff availability, meant that only a small sample of staff could be interviewed. The number of participants in a qualitative interview can affect the likelihood that saturation of the data is achieved. Fusch and Ness (2015) discuss how the decision of when saturation is achieved can be difficult, and that it depends not just on the thickness of the data (number of participants) but also on the richness of the data (depth, or quality). Bowen (2008) suggests that claims of saturation “should be supported by an explanation of how saturation was achieved and substantiated by clear evidence of its occurrence” (2008). In the present study, it is likely that saturation was achieved, because each of the six themes were endorsed with quotes from two to six participants. This demonstrates that no single, individual participant was generating important ideas that were unreplicated in interviews with other participants. The inclusion of staff from multiple roles in the sampling strategy ensured that the data had both a depth and a richness. Where similar points were raised by multiple participants, their differing roles meant they had different perspectives on the phenomena discussed.

The inclusion of views from other relevant healthcare staff, for example general practitioners, may have further enriched the themes developed. GPs are also implicated in the chain of referrals relating to mental health, for example in the referral of patients to IAPT services. Predictive technology to detect depression and anxiety could potentially be provided at a GP consultation, since GPs are likely to be aware of patients’ comorbidities that may predispose them to poor mental health. Any future research in the implementation of the predictive technology approach discussed here could therefore usefully involve consultation with GPs in addition to other community healthcare workers such as those included here.

### 8.4.5 Conclusions

This chapter has explored healthcare staff’s views of the implementation of technology for the prediction of depression and anxiety in older adults. Themes developed through the thematic analysis showed that while the integrated care team is a good potential host for such technology, and staff perceive benefits to its use, challenges exist around the engagement of older adults and staff with such a system. However, pathways to manage the care of depressed and anxious older adults already exist in this care context, and technology could be implemented within these, using the examples of Telehealth and Florence to build upon. This study provides a case study example of one setting where predictive technology could be introduced for patient benefit, so long as appropriate considerations are made to meet the challenges discussed.

# Chapter 9 – Discussion

## 9.1 Introduction

This chapter presents a summary of the four research studies, considers the strengths and limitations of the work, examines its contribution to knowledge, discusses its implications for policy, practice and research, and discusses future directions that could be taken to progress the work further. Research studies are considered within the context of the narrative literature review and the two mapping reviews described in Chapters 2 and 3.

## 9.2 Summaries of the studies

The aim of this thesis was to explore the practical application of machine learning to self-reported mood data to predict later depression and anxiety in older adults, and explore its potential for use in community healthcare. This section presents summaries of the four empirical studies which were conducted to achieve this aim.

### Study 1

The first mapping review demonstrated that many approaches have already been taken to the prediction of future depression and, to a lesser extent, anxiety, using machine learning (e.g. Alam et al., 2015; Bian et al, 2017; Devi & Kumar, 2016; Jin et al., 2015). However, none of these studies addressed the prediction of future depression or anxiety in an exclusively older adult population. Furthermore, these approaches largely based their predictive models either on medical histories (Bian et al., 2017; Dabek & Caban, 2015; Jin et al., 2015; Kessler et al., 2016; Wardenaar et al., 2014), one-time self-report scales (Devi & Kumar, 2016; Rude et al., 2010; Schalinski et al., 2016) or on neurological measures (Foland-Ross et al., 2015; Long et al., 2014; Mourao-Miranda et al., 2012). There was a lack of studies using dynamic data collection techniques over multiple days to predict future conditions. Collecting data over multiple days is an approach which may help to overcome the biasing effect that temporary mood states can have on one-time measures (Barsky, 2002). Thus Study 1 involved exploring the application of machine learning to daily self-reported mood from older adults to predict depression. Both supervised and unsupervised machine learning techniques were applied to the data, to use mood scores collected over one week to predict depression status, as measured via the 15-item GDS at 9 weeks’ follow-up. Cluster analysis, an unsupervised learning technique, was first applied to the data, and this produced a model for predicting depression with high sensitivity, but low specificity.

Supervised learning techniques were then explored. The LASSO was chosen for use because it had previously been used with samples of a similar size (Brann et al., 2017; Wardenaar et al., 2014), and it was designed to reduce the problem of overfitting. The LASSO was used with logistic regression and cross-validation to develop two predictive models. These models used cut-offs of five points and three points on the GDS, since literature suggested that these were both valid cut-points (Sheikh & Yesavage, 1986; Arthur et al., 1999). An ROC analysis of the two models produced areas under the ROC curve of 0.88 and 0.80 respectively. These areas were similar to, or greater than, AUCs for other machine learning approaches to predicting future depression reported in the literature. For example, Dabek & Caban (2015) reported an AUC of 0.76 for their predictive model, while Jin et al. (2015) reported AUCs of 0.78-0.81, and Jimenez-Serrano et al. (2015) reported AUCs of 0.66-0.74. The results therefore suggested that the novel approach taken in the study had potential for the future prediction of depression in older people, if validated prospectively.

### Study 2

The narrative literature review presented in Chapter 2 highlighted the varied views of older adults on the use of technology for various purposes. For example, research has reported that older adults enjoy the convenience that technology can provide (Chen & Chan, 2013, Mitzner et al., 2010) although ‘old-fashioned’ ways of doing things are often preferred over new methods facilitated by technology (Peek et al, 2015), and some aspects of technology are seen by older adults to intrude on their daily lives (Greenhalgh et al, 2013). Given these varied and, at times contradictory, views represented in the literature, it was considered important to understand the views of older people with regard to use of technology to report on mood and manage mental health. The second mapping review in Chapter 3B highlighted a lack of research on older adults’ attitudes towards technology designed to support mental health. Just three studies were found which addressed such attitudes (Crabb et al., 2012; Pugh et al., 2014; Sauve et al., 2016), and these were found to have methodological issues biasing them towards reporting positive views.

Study 2 therefore aimed to establish a more representative account of older adults’ views of technology to self-report mood and support mental health. It was considered important to understand older adults’ usability needs for such technologies in order to develop a new tool for their daily use. Furthermore, understanding the motivators and barriers to using technology for these purposes was also considered to be important to inform the development of the tool.

Following a pilot study, which informed the methods, two groups of older adults took part in interactive sessions where they completed discussion activities, as well as interacting with various apps and devices and reacting to these. Methods were inspired by the COBALT methodology which uses technology interaction within research study sessions, and assumes a user-as-expert approach (Astell et al., 2016). Participants’ reactions and discussions were video- and audio- recorded and then transcribed. Transcripts were analysed using Template Analysis (King 1998; 2004; 2012).

Participants described being motivated to use technology to alleviate low mood and loneliness. This supported Cotten and colleagues’ finding (2013) that technology can be beneficial to reduce loneliness in older people. In addition, Study 2 recognised increased self-reliance as a motivator to the use of technologies to support mental health. Barriers to the use of technology for mood-reporting and supporting mental health included a preference for human contact over use of technology. Although a preference for human contact over technology has been reported in previous literature about healthcare in general (Fischer et al., 2014), the present work demonstrates that this barrier is also relevant to the more specific field of mental health. The study demonstrated too that older adults fear what might happen if depression or anxiety are predicted by technology. Fear around the consequences of triggering certain devices has been found in previous research in the context of emergency pendant alarms (Greenhalgh et al, 2013). However, the findings in Study 2 demonstrate that this fear extends to the use of technology to report on mental health, and suggests that this may be caused by a fear of being subjected to compulsory treatment. These findings could inform the materials provided alongside technology used for these purposes, to inform users that the technology does not replace human staff, and that detection of symptoms of poor mental health will not lead to automatic and instantaneous intervention. Participants also discussed how symptoms of poor mental health might affect engagement with technology among older adults. Further research is recommended to explore this phenomenon further.

Findings around usability included that text and buttons used in apps for older people to support their mental health should be large to account for diminished visual acuity and dexterity. This reflected previous work in which older adults reported a fear that technology could cause eye-strain (Chen & Chan, 2013). A balance between adequate screen size to present these features, and portability of the device to be used, was found to be important in Study 2, and this supported work by Vaportzis et al. (2017). This balance was largely determined by personal preference, so allowing a choice of platform was found to be important for motivating continued use. The findings also suggested that developers should avoid assumptions about users’ knowledge of jargon and functional elements of apps and websites, as also reported in (Eisma et al., 2004). Clear instructional language and precise explanations are therefore needed to facilitate good usability.

### Study 3

The mapping review in Chapter 3A found that only a minority of prior studies had evaluated a predictive model for depression or anxiety using a prospective study design (Kessler et al., 2016; Mourao-Miranda et al., 2012; Rude et al., 2010; Tortajada et al., 2009). However, testing the models developed in Study 1 in a prospective paradigm was deemed important to evaluate their potential. Study 3 therefore involved collection of new mood and depression data from a sample of older adults similar to the participants whose data were used in Study 1. The aim of the study was to prospectively test the trained and cross-validated models for predicting depression status developed in Study 1.

Participants answered six questions about their mood and appetite online every day for one week, using a device of their own choice and ownership. At baseline and nine weeks follow-up, participants completed a GDS test and the anxiety subscale of the HADS online. The two models developed in Study 1 for the prediction of depression status at follow-up were tested with the data collected in this study, and an ROC analysis was used to examine the performance of the two models. Areas under the curve demonstrated that the models both provided an above-chance level of prediction (0.69 for the model with a cut-off of five points on the GDS and 0.63 for the model using a cut-off of three). These results were comparable to the results of the prospective application of machine learning techniques to predict future depression in a study by Kessler et al. (2016), where AUCs of 0.63 to 0.71 were reported. The results of Study 3 therefore indicated that the approach had potential, although the non-characteristic shape of the ROC curve for one of the models indicated that collecting more data from a larger sample would be necessary to produce a better estimate of the ROC curve. This would provide a greater level of confidence that any given choice of operating threshold would achieve the sensitivity and specificity suggested by the curve.

### Study 4

The mapping review in Chapter 3A found that while many studies (18 out of 19) explored the future prediction of depression, only a minority (6 out of 19) reported attempts to predict future anxiety. Some of these predicted anxiety alongside other conditions (Alam et al., 2016; Bian et al., 2017; Dabek & Caban, 2015; Long et al., 2014; Mourao-Miranda et al., 2012), but only one study reported an attempt to produce a model predictive of anxiety alone (Devi & Kumar, 2016). Study 4 in this thesis sought to apply the LASSO technique to the mood data collected from older adults to determine the possibility of predicting future anxiety based on these inputs. Data from the six mood and appetite reports and the anxiety subscale of the HADS were analysed using the LASSO with logistic regression and cross-validation. The LASSO selects variables which usefully contribute to a model and shrinks the coefficients for those variables included to avoid overfitting. When applied to the mood and anxiety data, the LASSO produced a predictive model which only included a constant – none of the mood and appetite variables (happy, sad, tired, alert, relaxed, hungry) were included in the model. This suggested that none of the mood variables were predictive of whether a participant would experience anxiety nine weeks later. This contrasted with the results for the depression models, and this may have been because anxiety has a more sudden onset than depression (Finlay-Jones & Brown, 2009; Angst & Vollrath, 1991), or because the mood words used in the study were not sensitive to any prodrome that anxiety may have. Use of the same approach, but replacing the mood words rated by participants with words used in existing anxiety scales, (for example ‘restless’, ‘worried’, ‘nervous’), may produce a more effective model. Further work would be necessary to validate this hypothesis. Nevertheless, the work in Study 4 has added to the literature on prediction of future anxiety using a novel, dynamic approach to the collection of input data, while previous approaches have only reported use of neurological measures (Mourao-Miranda et al., 2012; Long et al., 2014), records from healthcare practices (Dabek & Caban, 2015; Bian et al., 2017), biosensor data (Alam et al., 2016) or one-time self-report scales (Devi & Kumar, 2016). The approach described here represents an approach which is less invasive and cheaper than taking neurological measures, and gives the user more control than simply accessing their medical data. The approach therefore warrants further exploration.

### Study 5

The final empirical study in the project considered the implementation of a machine learning tool for predicting depression and anxiety within community healthcare, and sought the opinions of staff on the implementation of technology for this purpose. Semi-structured interviews were conducted with seven members of community healthcare staff to this end. Interviews were recorded on an encrypted voice recorder and were transcribed verbatim. Transcripts were analysed using Thematic Analysis (Braun & Clarke, 2006).

Six main themes were developed from the data using thematic analysis. The first theme focussed on suitability of the community healthcare team as a setting for implementing the predictive technology. The second theme explored the existing pathways for the detection and management of depression and anxiety described by participants. The findings from these two themes demonstrated that the community healthcare team already had pathways to deal with early signs of mental health conditions, including the provision of low level support from occupational therapists, and signposting to mental health nurses and GPs where necessary. The existence of these pathways suggests that this service is likely to be suitable to support the introduction of predictive technology for depression and anxiety.

Theme 3 addressed precedents in the use of technology to support the healthcare of older adults in the community. Participants described how two existing technologies are already being used for older adult patients to report on physical healthcare measures from their homes: the ‘Telehealth’ system; and ‘Florence’ (see Section 8.3, Theme 3). Data from these systems are received by the healthcare team at a ‘Single Point of Access’, where staff monitor the data and provide alerts to relevant teams where necessary. An online service for the self-report of mood, as described in this thesis, could provide predictions of future mental health status to the Single Point of Access, from where alerts could be sent to relevant staff to provide early interventions.

Participants in Study 5 perceived benefits to the use of predictive technology, and these were discussed in Theme 4. These included the potential for saving staff time and money spent on crisis care, since early intervention was perceived as a way to reduce the occurrence of future crises. Other perceived benefits included empowering patients through giving them greater control of their healthcare, and reducing the feeling that they were a burden on the healthcare system. Prior research has recognised the potential for technology to empower patients through enabling them to manage their own healthcare (Mitzner et al, 2010), while it has the added benefit of reducing the feeling that they are a burden to healthcare staff (Peek et al, 2015). The present study highlights the fact that healthcare staff recognise these benefits, and can understand the benefits of prevention, which have been advocated in the Five Year Forward View for Mental Health (Mental Health Taskforce, 2016).

The study also highlighted several challenges to the implementation of such technology with an older adult population in Themes 5 and 6. Practical challenges perceived by participants included the lack of wifi and mobile phone signal in older adults’ homes. Participants also cited the attitudes of staff as potentially problematic, for example they suggested that staff may avoid taking measures that might cause them to have extra work to do. Thus staff culture in community healthcare services was considered another issue that might prevent successful implementation. This finding supports prior work which has reported that staff culture in healthcare settings is a crucial determinant of the successful implementation of new technologies (Weger, 2012; Ackerman et al., 2012).

## 9.3 Overall summary

An approach was sought to address the low rates of diagnosis of depression and anxiety in older adults (Rodda et al, 2011; Allan et al, 2014). Daily monitoring and prediction were seen as useful approaches, because many cases of depression and anxiety in older people go undetected (Manthorpe & Illiffe, 2005), and aside from the distress and economic burden associated with these conditions, each of these has been associated with a higher rate of attempted suicide in older people (O’Connell et al., 2004; Chartrand et al., 2012). The approach explored in this thesis used technology to allow older adults to report on their mood, and applied machine learning techniques to attempt to predict depression and anxiety from these mood data. Alongside the machine learning work, the views of older adults were sought to inform the development of a new approach to mood reporting. From the outset, the potential application of this approach within healthcare pathways was also considered, and interviews were conducted with community healthcare staff to explore this. Interviews sought to facilitate an understanding of existing pathways for the management of depression and anxiety in older adults, as well as to explore how technology is currently used by these staff, and thus to inform how the approach could be used within current services.

Findings from these studies demonstrated the potential of the LASSO with logistic regression as a method for deriving predictive models of future depression in older adults. They demonstrate too the ability to collect daily mood reports from older adults using digital technologies of their own choice and ownership. Interactive group sessions with older adults informed the development of the approach, and these sessions showed how approaches to mood reporting can be made user-friendly for this population. Participants highlighted the importance of balancing simplicity and functionality, and the need to avoid assumptions around commonly used jargon and symbolism in app design (e.g. a cross symbol indicating the ability to close a window). Interviews with healthcare staff provided examples of how technology is currently employed with a similar approach for the self-report of physical measures in community healthcare in Sheffield. The interviews also demonstrated that referral pathways already exist in one healthcare trust in Sheffield for the management of early signs of depression and anxiety, and thus there is potential for predictive technology to be applied within existing pathways in this setting. Results from Study 3 showed that the six mood measures recorded were not predictive of later anxiety, although the daily mood report approach may still be valid for the prediction of this condition, if different mood words more related to the experience of anxiety were used.

## 9.4 Strengths and Limitations

**Strengths**

The work presented in this thesis has a number of strengths. It explores a novel approach to the problem of underdiagnosis and late detection of depression and anxiety in older people. This approach could provide improved chances to intervene early in the course of these conditions. Chapter 2 discussed how the stigma surrounding mental health conditions may limit the help-seeking behaviours of older adults experiencing these conditions. The approach described in this thesis, whereby older adults answer questions about their mood on a digital device, may help to address this barrier, since it is known that people experiencing mental health conditions are more likely to disclose symptoms when they believe a computer, and not a human, is analysing their symptoms (Lucas et al., 2014).

The LASSO is a well-established method (Tibshirani, 1996) which has been applied in previous research exploring the prediction of depression (Brann et al., 2017; Wardenaar et al., 2014), although it has not been applied to the prediction of future depression or anxiety in older people in particular. Applying cross-validation within the method helped to prevent overspecialisation of the models, and the use of Monte Carlo repetitions prevented favourable allocations of the data within the cross-validation. The outcome measure for depression, the GDS, is a gold standard measure, which was developed specifically for use with an older adult population, where symptomatology may differ from that of younger adults (Yesavage et al., 1983). When analysed using ROC analysis, the two models developed in Study 1 resulted in AUCs which compared well against those reported for other machine learning approaches applied to prediction of depression and anxiety in the literature. The models were found to predict future depressed status even in individuals who at baseline had scored below the cut-off for depression.

Another strength of the work was the involvement of older adults to consider practical aspects of applying the models developed within digital technologies. McCurdie and colleagues (2012) have argued for the involvement of end users in the design process for mHealth solutions. In this thesis, it was essential to conduct qualitative work with older adults to inform aspects such as the usability of mood-reporting tools, and to explore what motivates and demotivates the use of such tools. Findings from Study 2 informed the data collection procedure for Study 3, which is likely to have improved the usability of the research tools for participants.

The use of COBALT methodology in Study 2 also represents a strength of the thesis. The ‘user-as-expert’ approach used within this methodology (Astell et al. 2012; 2016) ensured that older adult participants felt comfortable and confident expressing their opinions. Interactive activities with apps and technologies allowed for participants’ immediate reactions to be captured, meaning the data were more likely to contain participants’ unfiltered views.

This project also benefits from a study to consider the implementation of the approach in healthcare pathways. Many developers of internet interventions struggle to achieve implementation of the solutions they develop. Vis and colleagues (2015) report that, while empirical evidence for digital mental health interventions grows, uptake into services remains slow. Ackerman et al. (2012) have discussed how “the design and deployment of new IT projects in complex medical settings would benefit from empirically informed understandings of, and responses to, the contingent properties of human-technology relations.“ Thus Study 5 provides strength to the project by providing an understanding of the factors relating to implementation of new approaches which are important to healthcare staff.

The recruitment strategy used in Study 5 ensured that participants represented a cross-section of staff from different levels and job roles. This meant that the results included the views of people working directly with patients as well as those who manage services. Multiple studies have demonstrated that involving staff at multiple levels can improve the chances that it will be successfully implemented (Weger et al. 2012; Ackerman et al. 2012; Kyratsis et al. 2012). It was therefore considered important that Study 5 recruited from multiple roles, and the analysis in Study 5 has benefited from this. For example, the interviews with service managers provided detail on the organisation of referral pathways, while interviews with nurses and an occupational therapist provided examples illustrating how pathways were experienced by patients, and how frontline staff themselves viewed their workloads.

The research presented in this thesis also supports a number of key policy recommendations. The NHS Five Year Forward View for Mental Health (NHS Mental Health Task Force, 2016) describes key aspects of how mental health should be managed and describes necessary changes for the improvement of services. The document sets out a vision for the future, which includes the promotion of good mental health, and targeting at-risk groups with evidence-based interventions. The approach to the prediction of future depression and anxiety presented in this thesis is in line with this vision, since it advocates an approach enabling early intervention.

In 2016, The Academy of Medical Sciences published its report on ‘Improving the health of the public by 2040’. The report describes how a new approach to public health is necessary in order to improve the health of the nation. According to the report, this new, ‘health of the public’ approach requires several important developments. One of these is the “harnessing of new technologies and the digital revolution” (p5). This development includes the use of innovative technologies to “help us achieve the shift towards prevention and early intervention” (p5). This thesis has described how technology could be harnessed with an older adult population to allow early intervention in depression and anxiety, in line with this development. The report also argues for transdisciplinary research, and for “strengthening and developing methods of engagement between researchers and the public” (p100). The work that contributed to this thesis has crossed the boundaries of traditional academic disciplines, involving machine learning techniques as well as qualitative interview and groupwork techniques. It has also involved members of the public, both as research participants and as attendees to two public engagement events focussed on the use of technology to support mental health in older adults. The research is thus timely and fits well with the aims set out by the Academy.

This work was carried out in Sheffield, South Yorkshire. The clinical commissioning group for South and Mid Yorkshire, Bassetlaw and North Derbyshire (SMYBNDCCG) has published a Sustainability and Transformation Plan covering the years 2016-2021 (SMYBNDCCG, no year). The plan includes an intention to put “prevention at the heart of what we do” and states an aim to increase investment in preventative strategies. Development of an intervention to predict future mental health conditions in older people thus corresponds well with healthcare aims in this region.

**Limitations**

There are also several limitations to the work. In Study 3, despite collecting data to test the model at the cut-off of five points on the GDS, there were too few participants scoring above this cut-off at follow-up to achieve the ratio of depressed to non-depressed participants assumed in the sample size calculation. This may have resulted from a recruitment bias, both in the source of participants and the inclusion criteria which were used. Participants were unpaid volunteers, and thus to take part in the study required them to be motivated to do so of their own accord. One symptom of depression is reduced motivation (American Psychological Association, 2013), so it may be that people experiencing the prodrome of depression did not volunteer for the study because they were affected by low motivation. Since participants in the NANA study were paid, and a greater proportion of participants in that study were found to score above the cut-off of five on the GDS, it may be that paying participants can help to overcome the low motivation symptomatic of the prodrome of depression. Thus providing payment to participants may be a useful way to increase engagement with less motivated older adults, who may be at greater risk of depression.

The low rate of depression at follow-up in Study 3 meant that the ROC curve had a low level of detail and was not a characteristic shape. Collecting data from a larger sample, or from a sample with a higher proportion of individuals scoring above five on the GDS at follow-up, would improve the likelihood of obtaining a characteristic shape on the ROC curve, permitting a better analysis of the performance of the model. The fact that both this model, and the model using a cut-off of three points, both provided an above-chance level of prediction when the models were applied prospectively indicates that this approach may provide a useful model if applied to a larger dataset.

Another limitation of the project was the failure to include individual effects within the machine learning models. Thus, the models presented here assume older adult mood patterns to be homogeneous, although it is known that a variety of factors affect mood at an individual level (Golder & Macy, 2011). The models presented in this thesis could thus be strengthened by adding the ability to take account of individual effects. While techniques exist to do this, such techniques require a larger sample size than that available in the original dataset presented here. Maas and Hox suggest the requirement for data from a minimum of 50 individuals to conduct such techniques (2005). Future studies could explore the application of multi-level modelling to the problem of predicting depression in older adults, if a dataset with an adequate sample size were available.

The qualitative studies in this project also had limitations. In Study 2, the use of group sessions may have increased the likelihood of socially desirable responding. Therefore, a set of one-to-one interviews alongside the group activities may have increased the content validity of the study. Furthermore, participants had higher levels of education than the national average, and a greater number of females were recruited than males. These factors resulted from the use of convenience sampling - recruiting only those individuals showing an interest in a study about technology may have biased the results by over-representing people with greater experience using digital technologies. However, the transcripts from the study contained a range of both positive and negative comments regarding the use of digital technologies, and participants described having a wide variety of difficulties in using technologies, which they perceived to be non-user-friendly. The findings were therefore useful for the development of user-friendly methods for mood-reporting for older adults, irrespective of these limitations.

Furthermore, the choice to recruit participants from an over-50s community group may have caused an under-representation of socially isolated older people in the sample, since all of those recruited were members of a social group. Thus, views on the use of technology to alleviate loneliness discussed in Chapter 5 may not reflect the views of older adults living more isolated lives. Mody et al (2008) have outlined considerations, principles and techniques for the successful recruitment of isolated older adults to research studies. Future work to explore the views of older adults on mental health technologies could consider approaches discussed in their work in order to obtain a more representative sample.

In addition, the sample size of 15 participants may be considered small. Sample size is relevant in considerations of whether a study reaches ‘saturation’. The concept of saturation is unclear and hard to measure – Fusch & Ness (2015) recognise the importance of both breadth (number of participants) and depth (time participants spend in the study, quality of the data obtained) within the concept of saturation. In Study 2, although the total number of participants was small, each group took part in interactive activities over a total period of four hours, meaning participants had much time to express their thoughts. Thus, the study could be considered to have depth in these terms, despite having a small number of participants.

In Study 5, while sampling from a wide range of roles and levels increased the diversity of experience, this resulted in only a small number of individuals being recruited from each job role, which may have led to bias in the findings. If the approach to predicting depression and anxiety described in this thesis were to be tested in a larger, clinical population, further research in the form of a process evaluation could usefully be conducted to explore the views of a larger number of frontline staff on the implementation of the approach.

One further limitation to Study 5 is that the study only recruited from one healthcare trust, limiting the generalisability of findings. However, the trust had recently undergone an organisational change involving an amalgamation of three community services which had previously been managed separately. This had implications for the organisation of referral pathways, and for the practices carried out by frontline staff. This organisational change had also been carried out across 24 other ‘pioneer’ healthcare trusts, and it is therefore likely that pathways and practices have become more uniform across these different trusts as a result. Therefore, it is likely that findings relating to the healthcare pathways discovered in this study may also be relevant to the other 24 pioneer sites where these changes have taken place. Should the changes undertaken in these areas be rolled out more widely, there is likely to be even greater applicability of the research findings to other settings.

## 9.5 Reflections on the use of the medical model

This thesis has argued for a new approach to the prediction of depression and anxiety in older adults with a view to enabling early intervention in the course of these conditions. The rationale for this includes cost-saving for health and social care services, where depression and anxiety in later life are correlated with increased overall cost of care (Unutzer et al., 2009; Adams et al, 2015). The thesis has taken a medical model approach, including a study to consider the implementation of digital technology for the prediction of depression and anxiety in healthcare services. However, this is not the only approach that could be taken with the predictive machine learning models developed in Studies 1 and 3. It would also be possible to use the models in a self-management paradigm, wherein alerts generated by the prediction of depression or anxiety could be used to prompt the user to take action to alleviate symptoms.

Some existing apps and websites take a self-management approach to depression and anxiety. These include ‘MoodMission’, an app which recommends activities that have an evidence base for relieving symptoms of both depression and anxiety. Other examples of apps offering evidence-based suggestions to maintain good mental health include ‘Pacifica’, ‘Calm’, and ‘Five Ways to Wellbeing’. While some of these apps also offer the ability to rate mood on a daily basis, no apps or websites currently use a machine learning approach to predict future outcomes of depression or anxiety in older adults, and make suggestions relevant specifically for this group. Thus the approach described in this thesis could be applied with suggestions for ways in which the user could improve their mental health, to break new ground in this area.

Implementing the models developed in this thesis in a self-management approach could have the benefit of further reducing costs to the NHS. Study 5 in this thesis (Chapter 8) showed that interventions currently used by community healthcare staff to address depression and anxiety in older adults include signposting to self-help literature and psychoeducation. These ways of addressing the management of mental health could be provided through digital technologies. For example, the webpages of the NHS offer psychoeducation on both depression and anxiety, as well as offering mp3 downloads that aid relaxation. Combining this type of approach with alerts from predictive models based on self-reported mood may offer a package which would allow some users to effectively self-manage their mental health before crises occur. Further research would be required to understand the benefits and disadvantages of this approach.

The target user group of the approach described in this thesis is older adults. Any self-management tools to be implemented with the models developed in this thesis should therefore be appropriate and acceptable to this target demographic. Study 2 explored the motivators, barriers and usability needs of older adults in the use of digital technologies to support mental health (Chapter 5). Participants in this study described ways in which they used digital technologies to change how they were feeling, through choosing a particular type of music to listen to, and through playing various games on tablet computers. These activities could be included within an app or website to help older people alleviate the early symptoms of depression or anxiety. The effectiveness and acceptability of including such activities within a digital self-management approach for older adults could now usefully be explored.

## 9.6 Implications of the research

### Implications for policy

This thesis has implications for both national and local policy. The Five Year Forward View for Mental Health strategy (Mental Health Taskforce, 2016) highlights the need for better crisis care and the need for better prevention of mental health conditions through public health interventions. However, the strategy does not describe measures to be taken where people are known to be at increased risk for mental health conditions. This is in spite of known predictors and correlates of mental health conditions, including chronic diseases and stressful life events (Easton et al. 2015; Roy & Lloyd, 2012; Cole & Dendukuri, 2003). This thesis has demonstrated an approach that could be used to monitor mood and predict future depression status in older people at increased risk. National policy on mental health could recommend that approaches should be taken to monitor mood in people at increased risk of depression and anxiety to identify early signs of declining mental health. Such an approach could allow early intervention to prevent later crises from occurring.

While the local Sustainability and Transformation Plan mentioned in Section 9.4 describes the importance of managing the mental health of children aged 0-15, it does not make any recommendations for the management of mental health in older adults. This thesis has highlighted the underdiagnosis of mental health conditions in older adults (Chapter 2) and explored older adults’ views of mental health and technology, finding that stigma around mental health conditions is still common among older adults (Chapter 5), and that screening for mental health conditions in community care in Sheffield is in decline (Chapter 8). While these are not major findings of the thesis, they nonetheless highlight the need for policy to address the mental health of older adults as a priority, and to support the provision of interventions that could help to detect the onset of poor mental health. Detecting depression and anxiety in older people is known to be difficult (Mental Health Taskforce, 2016), and older people are more prone to chronic diseases than the general population, as well as being more likely to suffer bereavement. These factors are known to increase the risk of depression and anxiety (Easton et al. 2015; Roy & Lloyd, 2012; Cole & Dendukuri, 2003). It is thus of importance that the mental health of older adults is considered in policy, and that measures are taken to manage the increased risk that certain factors of later life can cause.

### Implications for healthcare practice

This thesis has also provided evidence for ways that healthcare practice could be improved. Community healthcare in Sheffield is decreasing the use of screening practices for mental health conditions in older adults due to increased demand for services, lack of staff time, and cuts to services (Chapter 8). Employing digital technologies may offer a partial solution to this problem. Study 3 demonstrated that older adults are willing and able to use their own self-purchased digital devices to report on their mood (Chapter 6). Once further work is completed to ensure the effectiveness and cost-effectiveness of the machine learning approach, older adults in community care could be assessed online, at distance, with a monitoring approach similar to that used with Telehealth and Florence (Chapter 8). Thus while there are no instantaneous implications for practice as a result of this thesis, the groundwork is in place for future development of the approach, which could have a positive effect on the detection of depression and anxiety in older adults.

Findings in this thesis also have implications for the implementation of healthcare technologies with older adults more generally. Study 2 highlighted the fact that some older adults fear using healthcare technologies because they are unsure what might happen to them if the technology indicates there is a problem (Section 5.6 – Barriers). Participants in Study 2 also expressed a preference for human contact over solely interacting with technology (Section 5.6 – Barriers). These findings could inform how any new technology to be used with older adults in healthcare practice is implemented. They indicate that older adults would benefit from clear explanations from healthcare staff about how to use technologies, what their purpose is, and what might happen to them as a result of any information being shared using such technology. Explanations should emphasise too that while healthcare technologies may inform healthcare staff’s decision-making, the technologies themselves do not take decisions, or replace the care of (human) health professionals. The findings presented in this thesis suggest that such explanations and reassurances are likely to increase older adults’ confidence in using healthcare technologies.

### Implications for research

The thesis also has implications for research practice. Although the mapping review in Chapter 3 demonstrated that many machine learning approaches have been used in the prediction of mental health conditions, none addressed the prediction of future mental health conditions in older adults, and none had attempted to use dynamic data collection techniques for this purpose. Thus, this thesis has shown that machine learning approaches can be usefully applied to data collected dynamically over a period of one week, reducing the bias that results from using one-day, ‘snapshot’ approaches.

The thesis has also shown that research can be conducted with older adults using digital devices of their own preference and their own purchase to self-report on their mood (Chapter 6). Thus, it is not always necessary to provide older adults with digital technologies to collect data from them remotely in a research study. This knowledge could reduce costs for researchers interested in collecting self-report data from older adults, although researchers should be aware of any potential bias this may introduce to recruitment methods.

Furthermore, Study 2 demonstrated that the COBALT approach to interactive group sessions (Astell et al., 2016; 2012) could elicit important reactions and insight from participants, that may have been more difficult to obtain using a traditional focus group approach (Chapter 5). The interactive elements of the sessions encouraged participants to react immediately, ask for help when necessary, and discuss their experiences, all of which provided rich data to analyse. Thus, the work presented here provides support for the COBALT approach as an effective qualitative research methodology when working with older adults to develop interventions using digital technology.

## 9.7 Future directions

Studies 1 and 3 have provided evidence that supports the use of machine learning with self-reported mood data to predict future depression. Study 3 involved a prospective validation of the models developed using cut-offs of three and five points on the GDS in Study 1. For the model using a cut-off of three points on the GDS, it was possible to obtain a clear estimate of the ROC curve, and this showed the predictive ability of this model to be above chance-level, indicating promise for the approach. The predictive ability of the model developed at a cut-off of five was less clear, as it was not possible to achieve a detailed estimate of the ROC curve. This meant that confidence in the reported area under the curve for this model was low. This occurred as a result of a lower than expected incidence of depression in the prospective sample at follow-up, perhaps caused by unintended bias in the recruitment strategy (see Section 6.2 – Study 3). Future research could therefore explore the prospective validation of the model developed with a cut-off of five points in a new sample which includes a greater number of depressed older adults.

It would also be appropriate to test the approach discussed in this thesis in a population of older adults at increased risk of depression, since this is the population where the approach could most usefully be applied in practice. As discussed in Chapter 2, long term conditions such as diabetes and COPD are correlated with an increased risk of depression (see Section 2.3), and so testing the approach in a population of older adults with these conditions would be useful. The study populations in Study 1 and Study 3 were not recruited on the basis of their risk of depression, and indeed recruitment through the University of the Third Age in both studies was thought to artificially select for individuals at decreased risk of depression, as a result of their social engagement through this organisation (Section 6.4.1). Study 5 discussed the views of community healthcare staff and established some of the benefits and barriers to applying the approach in community healthcare pathways. Participants in Study 5 saw benefits to the use of the approach with older adults requiring in-home support for long term conditions. A large trial in this population could also include a process evaluation to understand the experience of both patients and healthcare staff.

Study 4 sought to explore the possibility of using the LASSO to develop a model to predict anxiety. Since the LASSO selected all six of the mood and appetite variables out of the model, it was not deemed possible to predict anxiety based on these measures. Future research could therefore explore using other words that may be more specific to the experience of anxiety, for example ‘restless’, ‘worried’, or ‘tense’. While it may be that the prodrome of anxiety is less consciously accessible than the prodrome of depression, further experiments using different mood words would indicate whether a subjective, self-report paradigm was a valid approach to its prediction.

For the prediction of depression and anxiety using digital technologies to be further developed, it would also be necessary to develop a standalone piece of software or app to enable effective and enjoyable user interaction, and effective dove-tailing with existing software used by healthcare professionals. Study 2 explored older adults’ usability requirements and preferences for interacting with digital technologies, and the findings from this study informed the design of the data collection tools for Study 3. Study 3 used Survey Monkey to collect the data from participants and this provided benefits as a research tool, including the ability to increase the font size of question text. However, it would not be possible to use Survey Monkey as the main data collection tool in an approach applied in general healthcare practice, since adherence to NHS security protocols would be required. Future work to develop a secure, NHS-based app or website could involve co-design with older adults to develop an interface which meets their needs and is motivating to use, while also assuring data protection. Furthermore, the co-design process could also involve healthcare professionals, from a variety of roles and levels, to inform the development of the staff-facing part of the system. Integration of such a tool with electronic patient registers, for example SystmOne or EMIS, would be beneficial, since this computer software is widely used in UK healthcare practice, and allows staff from different roles to share and access data held about a patient, when consent has been given by patients for such sharing to occur. Integrating self-reported mood scores into a central system such as SystmOne would allow historical mood data to be viewed by any member of staff with access to that system, which may help inform decisions taken by multiple healthcare staff.

It would also be important to ascertain the cost effectiveness of implementing the approach described. The aim of the approach overall is to enable effective targeting of early interventions to older adults predicted to become depressed or anxious, in order to avert the worsening of these conditions. NICE recommends a stepped model of care for depression and anxiety, where interventions become more intense, and thus more expensive, at higher steps of the model. The ability to intervene while patients are still at lower steps of the model thus has the potential to provide savings to the NHS, among other benefits. While the approach described in this thesis could help to identify those individuals most at risk, it would also cost money to implement. Thus a cost effectiveness study could help to elucidate what savings (if any) could be made through employing this system.

## 9.8 Contribution to knowledge

The original contribution to knowledge made by this thesis has several components:

1. The first systematic mapping review in Chapter 3A, found no studies applying machine learning techniques to data from older adults to predict future depression or anxiety. This thesis has proposed one method of predicting depression and anxiety in older adults, and has demonstrated the potential of this approach in the future prediction of depression. The application of machine learning, in particular the LASSO, to older adults’ self-reported mood and appetite data is a novel approach, and this work has now been published in a peer-reviewed journal.
2. The second mapping review in Chapter 3 showed that older adults’ views of using technology to support mental health are largely unknown. Chapter 5 describes a study conducted with older adults to understand their views of technology to monitor and support mental health, and to establish their usability requirements for mood reporting on digital devices. While this study supported findings in previous work examining older adults’ views of technology more generally, it also revealed a number of motivators and barriers that were new and specific to the use of technology to monitor and support mental health.
3. Previous work has demonstrated the importance of gaining input from staff at all levels of an organisation when new technologies are introduced. This thesis has recognised the importance of staff contributions, and has therefore featured interviews with NHS community healthcare staff, to gain an understanding of their views of implementing predictive technology with older adults. The knowledge generated from these interviews is vital for informing the implementation of the technology described here, and also informs the implementation of technology in healthcare more generally.
4. The knowledge generated through each of the studies in this thesis is complementary to the other studies, and as such, the work presented here is ‘more than the sum of its parts’. The machine learning approach, while being of interest in its own right, is of greater use when complemented by the insight from older adults as to how data capture can be presented most effectively. It is also more useful when considered alongside the views of NHS staff regarding barriers to its implementation. Thus while each study contributes new knowledge, the contribution to knowledge of this thesis is greatest when considered as a whole, as many existing approaches to machine learning presented in the literature fail to consider user interaction or implementation in their approach.

## 9.9 Conclusions

This thesis has built on previous work to apply machine learning techniques to the prediction of mental health conditions. It has focussed on the older adult population, and has considered how best to utilise the opportunities provided by digital technologies to predict future depression and anxiety in this population. It has considered the application of this approach from the perspectives of both older adult users and healthcare staff. The results have shown that it is possible to develop a model to predict future depression in older adults from their self-reported mood scores over one week. Furthermore, this model has been applied prospectively to estimate future depression status with an above-chance level of prediction. Older adults are able and willing to use technology to report on their mood, although some barriers exist to use. Motivators to the use of such technology include the ability to increase self-reliance. While many apps assume transferable knowledge in operating mechanisms, the research presented here has indicated that no such assumptions of prior knowledge should be made when designing for older adults.

Healthcare staff recognise the benefits of the approach described here, but perceive barriers to its implementation, including aspects of staff engagement, and the cost and supply of digital technologies. However, the fact that staff already work with remote technologies in physical healthcare, and that pathways exist for managing early signs of depression and anxiety, suggest that new approaches to the monitoring of older adults’ mental health using technology could be implemented in this setting. While further work is necessary to augment the models developed so far, and to further explore the possibility of predicting anxiety, the studies presented in this thesis together indicate that the approach is viable and worthy of further investigation. The work has provided a novel contribution to knowledge, and has implications more broadly for the field of technology usage with older adults.

# References

Ackerman, S. L., Tebb, K., Stein, J. C., Frazee, B. W., Hendey, G. W., Schmidt, L. A., & Gonzales, R. (2012). Benefit or burden? A sociotechnical analysis of diagnostic computer kiosks in four California hospital emergency departments. *Social science & medicine (1982), 75*, 2378-85. doi:10.1016/j.socscimed.2012.09.013

Adams, L. Y., Koop, P., Quan, H., & Norris, C. (2015). A population-based comparison of the use of acute healthcare services by older adults with and without mental illness diagnoses. *Journal of psychiatric and mental health nursing, 22*, 39-46. doi:10.1111/jpm.12169

Alam, M. G., Abedin, S. F., Ameen, M. A., & Hong, C. S. (2016). Web of objects based ambient assisted living framework for emergency psychiatric state prediction. *Sensors (Switzerland), 16*. doi:10.3390/s16091431

Aldarwish, M. M., & Ahmad, H. F. (2017). Predicting Depression Levels Using Social Media Posts. *2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS)*, (pp. 277-280). doi:10.1109/ISADS.2017.41

Aldrich, J. (1998). Doing Least Squares: Perspectives from Gauss and Yule. *International Statistical Review, 66*, 61-81. doi:10.1111/j.1751-5823.1998.tb00406.x

Allan, C. E., Valkanova, V., & Ebmeier, K. P. (2014). Depression in older people is underdiagnosed. *The Practitioner, 258*, 19-22.

Almeida, O. P., & Almeida, S. A. (1999). Short versions of the geriatric depression scale: a study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. *International Journal of Geriatric Psychiatry, 14*, 858-865. doi:10.1002/(SICI)1099-1166(199910)14:10<858::AID-GPS35>3.0.CO;2-8

Alpass, F. M., & Neville, S. (2003). Loneliness, health and depression in older males. *Aging & Mental Health, 7*, 212-216. doi:10.1080/1360786031000101193

Alpaydin, E. (2004). *Introduction to machine learning.* Cambrige, MA and London, UK: MIT Press

Andrews, G., Cuijpers, P., Craske, M. G., McEvoy, P., & Titov, N. (2010). Computer Therapy for the Anxiety and Depressive Disorders Is Effective, Acceptable and Practical Health Care: A Meta-Analysis. (B. T. Baune, Ed.) *PLoS ONE, 5*, e13196. doi:10.1371/journal.pone.0013196

Andrews, J. A., Harrison, R. F., Brown, L. J., MacLean, L. M., Hwang, F., Smith, T., . . . Astell, A. J. (2017). Using the NANA toolkit at home to predict older adults' future depression. *Journal of Affective Disorders*. doi:10.1016/j.jad.2017.02.019

Angst, J., & Vollrath, M. (1991). The natural history of anxiety disorders. *Acta Psychiatrica Scandinavica, 84*, 446-452. doi:10.1111/j.1600-0447.1991.tb03176.x

American Psychological Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders 5.* doi:10.1176/appi.books.9780890425596.744053

Appleby, J., Galea, A., & Murray, R. (2014). *The NHS productivity challenge: experience from the front line.* Technical report.

Arthur, A., Jagger, C., Lindesay, J., Graham, C., & Clarke, M. (1999). Using an annual over-75 health check to screen for depression: validation of the short geriatric depression scale (GDS15) within general practice. *International Journal of Geriatric Psychiatry, 14*, 431-439. doi:10.1002/(SICI)1099-1166(199906)14:6<431::AID-GPS937>3.0.CO;2-I

Astell, A. (2013). Technology and Fun for a Happy Old Age. In *Technologies for Active Aging* (pp. 169-187). Boston, MA: Springer US. doi:10.1007/978-1-4419-8348-0\_10

Astell, A. J., Hwang, F., Brown, L. J., Timon, C., Maclean, L. M., Smith, T., Adlam, T., Khadra, H., & Williams, E. (2014). Validation of the NANA (Novel Assessment of Nutrition and Ageing) touch screen system for use at home by older adults. *Experimental Gerontology, 60*, 100-107. doi:10.1016/j.exger.2014.10.008

Astell, A. J., Hwang, F., Williams, E. A., Archer, L., Harney-Levine, S., Wright, D., & Ellis, M. (2016). Well-Being and HCI in Later Life - What Matters? Springer International Publishing. doi:10.1007/978-3-319-40397-7\_42

Büchtemann, D., Luppa, M., Bramesfeld, A., Riedel-Heller, S. (2012). Incidence of late-life depression: A systematic review. *Journal of Affective Disorders, 142*, 172-179. doi:10.1016/j.jad.2012.05.010

Barsky, A. J. (2002). Forgetting, Fabricating, and Telescoping. *Archives of Internal Medicine, 162,* 981-984.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of general psychiatry, 4*, 561-71.

Bell, M. L., Fairclough, D. L., Fiero, M. H., & Butow, P. N. (2016). Handling missing items in the Hospital Anxiety and Depression Scale (HADS): a simulation study. *BMC research notes, 9*, 479. doi:10.1186/s13104-016-2284-z

Berkman, L. I., Berkman, C. A., Stanislav, K. A., Freeman, D. A., Linda, L. E., Ostfeld, A. D., Cornoni-Huntley, J. & Brody, J. A. (1986). Depressive symptoms in relation to physical health and functioning in the elderly. *American Journal of Epidemiology, 124*, 372-388. doi:10.1093/oxfordjournals.aje.a114408

Bhar, S. S., & Brown, G. K. (2012). Treatment of Depression and Suicide in Older Adults. *Cognitive and Behavioral Practice, 19*, 116-125. doi:10.1016/j.cbpra.2010.12.005

Bian, J., Barnes, L. E., Chen, G., & Xiong, H. (2017). Early Detection of Diseases using Electronic Health Records Data and Covariance-Regularized Linear Discriminant Analysis. *2017 Ieee Embs International Conference on Biomedical & Health Informatics (Bhi)*, 457-460.

Boeker, M., Vach, W., & Motschall, E. (2013). Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough. *BMC Medical Research Methodology, 13*, 131. doi:10.1186/1471-2288-13-131

Booth, A., Papaioannou, D., & Sutton, A. (2016). *Systematic approaches to a successful literature review* (2nd ed.). London: Sage Publications.

Bourland, S. L., Stanley, M. A., Snyder, A. G., Novy, D. M., Beck, J. G., Averill, P. M., & Swann, A. C. (2000). Quality of life in older adults with generalized anxiety disorder. *Aging & Mental Health, 4*, 315-323. doi:10.1080/713649960

Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: a research note. *Qualitative Research, 8*, 137-152. doi:10.1177/1468794107085301

Brann, E., Papadopoulos, F., Fransson, E., White, R., Edvinsson, A., Hellgren, C., Kamali-Moghaddam, M., Bostrom, A., Schioth, H.B., & Sundstrom-Poromaa, I. (2017). Inflammatory markers in late pregnancy in association with postpartum depression - A nested case-control study. *Psychoendoneurocrinology,* *79*, 146-159.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77-101. doi:10.1191/1478088706qp063oa

Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science, 16*, 199-215.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1993). *Classification and regression trees.* London: Chapman & Hall.

Brenes, G. A. (2007). Anxiety, depression, and quality of life in primary care patients. *Primary care companion to the Journal of clinical psychiatry, 9*, 437-43.

Brown, L. J., & Astell, A. (2012). Assessing mood in older adults: a conceptual review of methods and approaches. *International psychogeriatrics / IPA, 24*, 1197-206. doi:10.1017/S1041610212000075

Brown, L. J., Adlam, T., Hwang, F., Khadra, H., Maclean, L. M., Rudd, B., . . . Astell, A. J. (2016). Computerized Self-Administered Measures of Mood and Appetite for Older Adults. *Journal of Applied Gerontology*. doi:10.1177/0733464816630636

Bryant, C., Jackson, H., & Ames, D. (2008). The prevalence of anxiety in older adults: methodological issues and a review of the literature. *Journal of affective disorders, 109*, 233-50. doi:10.1016/j.jad.2007.11.008

Burns, M. N., Begale, M., Duffecy, J., Gergle, D., Karr, C. J., Giangrande, E., & Mohr, D. C. (2011). Harnessing context sensing to develop a mobile intervention for depression. *Journal of Medical Internet Research, 13*, e55. doi:10.2196/jmir.1838

Chartrand, H., Sareen, J., Toews, M., & Bolton, J. M. (2012). Suicide attempts versus nonsuicidal self-injury among individuals with anxiety disorders in a nationally representative sample. *Depression and Anxiety*. doi:10.1002/da.20882

Chen, K., & Chan, A. H.-S. (2013). Use or non-use of gerontechnology - a qualitative study. *International journal of environmental research and public health, 10*, 4645-66. doi:10.3390/ijerph10104645

Clark, R. A., Inglis, S. C., McAlister, F. A., F Cleland, J. G., & Stewart, S. (2007). Telemonitoring or structured telephone support programmes for patients with chronic heart failure: systematic review and meta-analysis. *British Medical Journal*. doi:10.1136/bmj.39156.536968.55

Cole, M. G., & Dendukuri, N. (2003). Risk Factors for Depression Among Elderly Community Subjects: A Systematic Review and Meta-Analysis. *American Journal of Psychiatry, 160*, 1147-1156. doi:10.1176/appi.ajp.160.6.1147

Comer, R. J. (2013). *Abnormal Psychology* (8th ed.). New York: Worth Publishers.

Conner, K. O., Lee, B., Mayers, V., Robinson, D., Reynolds, C. F., Albert, S., & Brown, C. (2010). Attitudes and beliefs about mental health among African American older adults suffering from depression. *Journal of Aging Studies, 24*, 266-277. doi:10.1016/j.jaging.2010.05.007

Conwell, Y., Duberstein, P. R., & Caine, E. D. (2002). Risk factors for suicide in later life. *Biological Psychiatry, 52*, 193-204. doi:10.1016/S0006-3223(02)01347-1

Cotten, S. R., Anderson, W. A., & McCullough, B. M. (2013). Impact of Internet use on loneliness and contact with others among older adults: Cross-sectional analysis. *Journal of Medical Internet Research, 15*, 215-227. doi:http://dx.doi.org/10.2196/jmir.2306

Cotton, R., Wilkins, A., & Young, C. (2014). The future's digital: Mental health and technology. Technical Report.

Coughlan, M. (2013). *Doing a literature review in nursing, health and social care.* London: Sage Publications.

Crabb, R. M., Rafie, S., & Weingardt, K. R. (2012). Health-Related Internet Use in Older Primary Care Patients. *Gerontology, 58*, 164-170. doi:10.1159/000329340

Crabb, R., & Hunsley, J. (2006). Utilization of mental health care services among older adults with depression. *Journal of Clinical Psychology, 62*, 299-312. doi:10.1002/jclp.20231

Cronin, P., Ryan, F., & Coughlan, M. (2008). Undertaking a literature review: a step-by-step approach. *British Journal of Nursing*, *17,* (1) 38-43.

Dabek, F., & Caban, J. J. (2015). A Neural Network Based Model for Predicting Psychological Conditions. In Y. Guo, K. Friston, A. Faisal, S. Hill, & H. Peng (Ed.), *International Conference on Brain Informatics and Health*, *9250*, pp. 252-261. doi:10.1007/978-3-319-23344-4\_25

Damián, J., Pastor-Barriuso, R., Valderrama-Gama, E., & Pedro-Cuesta, J. (2017). Association of detected depression and undetected depressive symptoms with long-term mortality in a cohort of institutionalised older people. *Epidemiology and Psychiatric Sciences, 26*, 189-198. doi:10.1017/S2045796015001171

Darkins, A. W., & Cary, M. A. (2000). *Telemedicine and telehealth : principles, policies, performances, and pitfalls.* New York: Springer

De Weger, E., MacInnes, D., Enser, J., Francis, S. J., & Jones, F. W. (2013). Implementing video conferencing in mental health practice. *Journal of Psychiatric and Mental Health Nursing, 20*, 448-454. doi:10.1111/j.1365-2850.2012.01947.x

Devi, S., & Kumar, S. (2016). An Adaptive Neuro Fuzzy Inference System for Prediction of Anxiety of Students. *8th International Conference on Advanced Computational Intelligence Chiang Mai, Thailand; February 14-16, 2016 An*, (pp. 7-13).

Djernes, J. K. (2006). Prevalence and predictors of depression in populations of elderly: a review. *Acta Psychiatrica Scandinavica, 113*, 372-387. doi:10.1111/j.1600-0447.2006.00770.x

Donker, T., Petrie, K., Proudfoot, J., Clarke, J., Birch, M.-R., & Christensen, H. (2013). Smartphones for smarter delivery of mental health programs: a systematic review. *Journal of Medical Internet Research, 15*, e247. doi:10.2196/jmir.2791

Drake, G., Csipke, E., & Wykes, T. (2013). Assessing your mood online: acceptability and use of Moodscope. *Psychological Medicine, 43*, 1455-64. doi:10.1017/S0033291712002280

Driscoll, H. C., Basinski, J., Mulsant, B. H., Butters, M. A., Dew, M. A., Houck, P. R., . . . Reynolds, C. F. (2005). Late-onset major depression: clinical and treatment-response variability. *International Journal of Geriatric Psychiatry, 20*, 661-667. doi:10.1002/gps.1334

Drozd, F., Vaskinn, L., Bergsund, H. B., Haga, S. M., Slinning, K., & Bjørkli, C. A. (2016). The implementation of internet interventions for depression: A scoping review. *Journal of Medical Internet Research, 18,* (9) e236 doi:10.2196/jmir.5670

Easton, K., Coventry, P., Lovell, K., Carter, L.-A., & Deaton, C. (2015). Prevalence and Measurement of Anxiety in Samples of Patients With Heart Failure: Meta-analysis. *Journal of Cardiovascular Nursing*, *31*, (4), 367-379 doi:10.1097/JCN.0000000000000265

Eisma, R., Dickinson, A., Goodman, J., Syme, A., Tiwari, L., & Newell, A. F. (2004). Early user involvement in the development of information technology-related products for older people. *Universal Access in the Information Society, 3*, 131-140. doi:10.1007/s10209-004-0092-z

Elliot, A. J., Mooney, C. J., Douthit, K. Z., & Lynch, M. F. (2014). Predictors of older adults' technology use and its relationship to depressive symptoms and well-being. *Journals of Gerontology: Series B, Psychological sciences and social sciences, 69*, 667-77. doi:10.1093/geronb/gbt109

Finlay-Jones, R., & Brown, G. W. (1981). Types of stressful life event and the onset of anxiety and depressive disorders. *Psychological Medicine, 11*, 803. doi:10.1017/S0033291700041301

First, M. B. (2016). *SCID-5-CV : Structured Clinical Interview For DSM-5 Disorders, Clinician Version.* American Psychiatric Association.

Fischer, S. H., David, D., Crotty, B. H., Dierks, M., & Safran, C. (2014). Acceptance and use of health information technology by community-dwelling elders. *International journal of medical informatics, 83*, 624-35. doi:10.1016/j.ijmedinf.2014.06.005

Fiske, A., Wetherell, J. L., & Gatz, M. (2009). Depression in older adults. *Annual review of clinical psychology, 5*, 363-89. doi:10.1146/annurev.clinpsy.032408.153621

Foland-Ross, L. C., Sacchet, M. D., Prasad, G., Gilbert, B., & Thompson, P. M. (2015). Cortical thickness predicts the first onset of major depression in adolescence. *46*, 125-131.

Fusch, P. I., & Ness, L. R. (2015). Are We There Yet? Data Saturation in Qualitative Research. *The Qualitative Report, 20*, 1408-1416.

Golder, S., & Macy, M. (2011). Diurnal and Seasonal Mood Vary with Work, Sleep and Daylength Across Diverse Cultures. *Science, 333*, 1878-1881. doi:10.1126/science.1202775

Greenhalgh, T., Wherton, J., Sugarhood, P., Hinder, S., & Procter, R. (2013). What matters to older people with assisted living needs? A phenomenological analysis of the use and non-use of telehealth and telecare. *Social Science & Medicine, 93*, 86-94.

Halfin, A. (2007). Depression: The benefits of early and appropriate treatment. *American Journal of Managed Care*. doi:6821 [pii]

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning.* New York: Springer. doi:10.1007/978-0-387-84858-7

Hu, Q., Li, A., Heng, F., Li, J., & Zhu, T. (2015). Predicting Depression of Social Media User on Different Observation Windows. *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 361-364. doi:10.1109/WI-IAT.2015.166

Jamshed, S. (2014). Qualitative research method - interviewing and observation. *Journal of Basic and Clinical Pharmacy, 5*, 87-8. doi:10.4103/0976-0105.141942

Jiménez-Serrano, S., Tortajada, S., & García-Gómez, J. M. (2015). A Mobile Health Application to Predict Postpartum Depression Based on Machine Learning. *Telemedicine Journal and e-Health, 21*, 567-74. doi:10.1089/tmj.2014.0113

Jin, H., Wu, S., & Di Capua, P. (2015). Development of a Clinical Forecasting Model to Predict Comorbid Depression Among Diabetes Patients and an Application in Depression Screening Policy Making. *Preventing chronic disease, 12*, E142.

Karstoft, K.-I., Galatzer-Levy, I. R., Statnikov, A., Li, Z., Shalev, A. Y., & members of the J-TOP group. (2015). Bridging a translational gap: using machine learning to improve the prediction of PTSD. *BMC psychiatry, 15*, 30.

Kessler, E.-M., Agines, S., & Bowen, C. E. (2014). Attitudes towards seeking mental health services among older adults: personal and contextual correlates. *Aging & Mental Health*, *19,* (2) 1-10. doi:10.1080/13607863.2014.920300

Kessler, R. C., van Loo, H. M., Wardenaar, K. J., Bossarte, R. M., Brenner, L. A., Cai, T., Ebert, D. D., Hwang, I., Li, J., de Jonge, P., Nierenberg, A. A., Metukova, M. V., Rossilini, A.J., Sampson, N. A., Schoevers, R. A., Wilcox, M. A., & Zaslavsky, A. M. (2016). Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. *Molecular Psychiatry, 21*, 1366-1371. doi:10.1038/mp.2015.198

King, N. (1998). Template Analysis. In G. Symon, & C. Cassell (Eds.), *Qualitative Methods and Analysis in Organizational Research* (pp. 118-134). London: Sage.

King, N. (2012). Doing template analysis. In G. Symon, & C. Cassell (Eds.), *Qualitative Organizational Research: Core methods and current challenges* (pp. 426-450). London: Sage.

King, N. (2004). Using templates in the thematic analysis of texts. In C. Cassell & G. Symon (Eds.), *Essential guide to qualitative methods in organizational research*, (pp. 256-270). London: Sage.

Kiosses, D. N., Szanto, K., & Alexopoulos, G. S. (2014). Suicide in older adults: the role of emotions and cognition. *Current psychiatry reports, 16*, 495. doi:10.1007/s11920-014-0495-3

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*, 606-13. doi:10.1046/j.1525-1497.2001.016009606.x

Krstajic, D., Buturovic, L. J., Leahy, D. E., & Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics, 6*, 10. doi:10.1186/1758-2946-6-10

Kupfer, D. J., Frank, E., & Perel, J. M. (1989). The Advantage of Early Treatment Intervention in Recurrent Depression. *Archives of General Psychiatry, 46*, 771. doi:10.1001/archpsyc.1989.01810090013002

Kyratsis, Y., Ahmad, R., & Holmes, A. (2012). Technology adoption and implementation in organisations: comparative case studies of 12 English NHS Trusts. *BMJ Open, 2*, e000872--e000872. doi:10.1136/bmjopen-2012-000872

Lam, C. L., Chin, W. Y., Lee, P. W., Lo, Y. Y., Fong, D. Y., & Lam, T. P. (2009). Unrecognised psychological problems impair quality of life and increase consultation rates in Chinese elderly patients. *International Journal of Geriatric Psychiatry, 24*, 979-89. doi:10.1002/gps.2210

Lasalvia, A., Zoppei, S., Van Bortel, T., Bonetto, C., Cristofalo, D., Wahlbeck, K., . . . Thornicroft, G. (2013). Global pattern of experienced and anticipated discrimination reported by people with major depressive disorder: a cross-sectional survey. *Lancet, 381*, 55-62. doi:10.1016/S0140-6736(12)61379-8

Lee, C., & Coughlin, J. F. (2015). PERSPECTIVE: Older Adults' Adoption of Technology: An Integrated Approach to Identifying Determinants and Barriers. *Journal of Product Innovation Management, 32*, 747-759. doi:10.1111/jpim.12176

Lee, H. J., & Dugan, E. (2015). How Large Is the Gap Between Self-Report and Assessed Mental Health and Does It Impact Older Adult Mental Health Service Utilization? *Journal of Gerontological Social Work, 58*, 3-19. doi:10.1080/01634372.2014.919978

Lesher, E. L., & Berryhill, J. S. (1994). Validation of the Geriatric Depression Scale - Short Form among inpatients. *Journal of Clinical Psychology, 50*, 256-60.

Libbrecht, M. W., & Noble, W. S. (2015). Machine learning applications in genetics and genomics. *Nature Reviews Genetics, 16*, 321-332. doi:10.1038/nrg3920

Likamwa, R., Liu, Y., Lane, N. D., & Zhong, L. (2013). MoodScope: Building a Mood Sensor from Smartphone Usage Patterns. *MobiSys '13 Proceeding of the 11th annual international conference on Mobile systems, applications and services*, 389-402. doi:10.1145/2462456.2464449

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing, 7*, 76-80. doi:10.1109/MIC.2003.1167344

Long, J., Huang, X., Liao, Y., Hu, X., Hu, J., Lui, S., . . . Gong, Q. (2014). Prediction of post-earthquake depressive and anxiety symptoms: a longitudinal resting-state fMRI study. *Scientific reports, 4*, 6423. doi:10.1038/srep06423

Lucas, G. M., Gratch, J., King, A., & Morency, L.-P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior, 37*, 94-100. doi:10.1016/J.CHB.2014.04.043

Luxton, D. D., McCann, R. a., Bush, N. E., Mishkind, M. C., & Reger, G. M. (2011). mHealth for mental health: Integrating smartphone technology in behavioral healthcare. *Professional Psychology: Research and Practice, 42*, 505-512. doi:10.1037/a0024485

Maas, C. J., & Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology, 1*, 86-92. doi:10.1027/1614-2241.1.3.86

Manthorpe, J., & Iliffe, S. (2006). *Depression in Later Life.* London & Philidelphia: JKP.

Marshall, C., & Rossman, G. B. (2006). *Designing qualitative research.* London: Sage

Martín-Merino, E., Ruigómez, A., Wallander, M.-A., Johansson, S., & García-Rodríguez, L. A. (2010). Prevalence, incidence, morbidity and treatment patterns in a cohort of patients diagnosed with anxiety in UK primary care. *Family Practice, 27*, 9-16. doi:10.1093/fampra/cmp071

Menaghan, E. G., & Lieberman, M. A. (1986). Changes in depression following divorce: A panel study. *Journal of Marriage and Family, 48*, 319-328. doi:10.2307/352399

Mental Health Taskforce. (2016). The Five Year Forward View For Mental Health. Technical Report.

Miloyan, B., Bulley, A., Bandeen-Roche, K., Eaton, W. W., & Gonçalves-Bradley, D. C. (2016). Anxiety disorders and all-cause mortality: systematic review and meta-analysis. *Social psychiatry and psychiatric epidemiology, 51*, 1467-1475. doi:10.1007/s00127-016-1284-6

Mitzner, T. L., Boron, J. B., Fausset, C. B., Adams, A. E., Charness, N., Czaja, S. J., . . . Sharit, J. (2010). Older Adults Talk Technology: Technology Usage and Attitudes. *Computers in Human Behavior, 26*, 1710-1721. doi:10.1016/j.chb.2010.06.020

Mody, L., Miller, D. K., McGloin, J. M., Freeman, M., Marcantonio, E. R., Magaziner, J., & Studenski, S. (2008). Recruitment and Retention of Older Adults in Aging Research. *Journal of the American Geriatrics Society, 56*, 2340-2348. doi:10.1111/j.1532-5415.2008.02015.x

Mohr, D. C., Lyon, A. R., Lattie, E. G., Reddy, M., & Schueller, S. M. (2017a). Accelerating Digital Mental Health Research From Early Design and Creation to Successful Implementation and Sustainment. *Journal of medical Internet research, 19*, e153. doi:10.2196/jmir.7725

Mohr, D. C., Zhang, M., & Schueller, S. M. (2017b). Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning. *Annual Review of Clinical Psychology, 13*, 23-47. doi:10.1146/annurev-clinpsy-032816-044949

Mourao-Miranda, J., Oliveira, L., Ladouceur, C. D., Marquand, A., Brammer, M., Birmaher, B., . . . Phillips, M. L. (2012). Pattern Recognition and Functional Neuroimaging Help to Discriminate Healthy Adolescents at Risk for Mood Disorders from Low Risk Adolescents. *PLoS ONE, 7*, e29482. doi:10.1371/journal.pone.0029482

National Institute for Health and Care Excellence (NICE). (2011). *Generalised anxiety disorder and panic disorder in adults: management. Clinical Guideline CG113.* NICE.

National Institute for Health and Care Excellence (NICE). (2016). *Depression in adults: recognition and management. Clinical Guideline CG90.* NICE.

Naylor, C., Parsonage, M., McDaid, D., Knapp, M., Fossey, M., & Galea, A. (2012). Long-term conditions and mental health: the cost of co-morbidities. doi:ISBN 9781857176339

Nordhus, I. H. (2008). Manifestations of depression and anxiety in older adults. In R. Woods, & L. Clare (Eds.), *Handbook of the Clinical Psychology of Ageing (Second Edition)* (2nd ed., pp. 97-110). Chichester: Wiley.

Obuchowski, N. (2005). ROC Analysis. *American Journal of Roentgenology, 184*, 364-372. doi:10.2214/ajr.184.2.01840364

O'Connell, H., Chin, A.-V., Cunningham, C., & Lawlor, B. A. (2004). Recent developments: suicide in older people. *BMJ (Clinical research ed.), 329*, 895-9. doi:10.1136/bmj.329.7471.895

Ofcom. (2015). *Adults' media use and attitudes report 2015.* Technical report.

Ooi, K. E., Lech, M., & Allen, N. B. (2013). Multichannel Weighted Speech Classification System for Prediction of Major Depression in Adolescents. *Ieee Transactions on Biomedical Engineering, 60*, 497-506. doi:10.1109/tbme.2012.2228646

Pachana, N. A., Byrne, G. J., Siddle, H., Koloski, N., Harley, E., & Arnold, E. (2007). Development and validation of the Geriatric Anxiety Inventory. *International Psychogeriatrics, 19*, 103. doi:10.1017/S1041610206003504

Paplos, K. G., Kontaxakis, V. P., Havaki-Kontaxaki, B. J., & Rabavilas, A. D. (2003). Suicidal intention, depression and anhedonia among suicide attempters. *Annals of General Hospital Psychiatry, 2*, S109. doi:10.1186/1475-2832-2-S1-S109

Passos, I. C., Mwangi, B., Cao, B., Hamilton, J. E., Wu, M.-J., Zhang, X. Y., . . . Soares, J. C. (2016). Identifying a clinical signature of suicidality among patients with mood disorders: A pilot study using a machine learning approach. *Journal of Affective Disorders, 193*, 109-116.

Peek, S. T., Luijkx, K. G., Rijnaard, M. D., Nieboer, M. E., Voort, C. S., Aarts, S., . . . Wouters, E. J. (2015). Older Adults' Reasons for Using Technology while Aging in Place. *Gerontology*. doi:10.1159/000430949

Pew Research Centre. (2014). *Older adults and technology use.* Technical report

Phillipson, L., Magee, C., Jones, S., Reis, S., & Skaldzien, E. (2015). Dementia attitudes and help-seeking intentions: an investigation of responses to two scenarios of an experience of the early signs of dementia. *Aging & Mental Health, 19*, 968-977. doi:10.1080/13607863.2014.995588

Poison, J., & Croy, S. (2013). Differentiating dementia, delirium and depression. *Nursing Times, 111*, 18-19.

Polisena, J., Tran, K., Cimon, K., Hutton, B., McGill, S., & Palmer, K. (2009). Home telehealth for diabetes management: a systematic review and meta-analysis. *Diabetes, Obesity and Metabolism, 11*, 913-930. doi:10.1111/j.1463-1326.2009.01057.x

Polisena, J., Tran, K., Cimon, K., Hutton, B., McGill, S., Palmer, K., & Scott, R. E. (2010). Home telehealth for chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Journal of Telemedicine and Telecare, 16*, 120-127. doi:10.1258/jtt.2009.090812

Préville, M., Mechakra Tahiri, S. D., Vasiliadis, H.-M., Quesnel, L., Gontijo-Guerra, S., Lamoureux-Lamarche, C., & Berbiche, D. (2015). Association between perceived social stigma against mental disorders and use of health services for psychological distress symptoms in the older adult population: validity of the STIG scale. *Aging & Mental Health, 19*, 464-74. doi:10.1080/13607863.2014.944092

Preschl, B., Wagner, B., Forstmeier, S., & Maercker, A. (2011). E-health interventions for depression, anxiety disorder, dementia, and other disorders in older adults: A review. *Journal of Cybertherapy and Rehabilitation* *4*, 371-385.

Pugh, N. E., Hadjistavropoulos, H. D., Klein, B., & Austin, D. W. (2014). A Case Study Illustrating Therapist-Assisted Internet Cognitive Behavior Therapy for Depression. *Cognitive and Behavioral Practice, 21*, 64-77.

Ramsey, A., Lord, S., Torrey, J., Marsch, L., & Lardiere, M. (2016). Paving the Way to Successful Implementation: Identifying Key Barriers to Use of Technology-Based Therapeutic Tools for Behavioral Health Care. *The Journal of Behavioral Health Services & Research, 43*, 54-70. doi:10.1007/s11414-014-9436-5

Rapoport, M., Seitz, D., Wiens, A., & Lily, E. (2016). *Geriatric psychiatry review and exam preparation guide : a case-based approach.* Toronto: University of Toronto Press.

Reynolds, C. F., Thomas, S. B., Morse, J. Q., Anderson, S. J., Albert, S., Dew, M. A., . . . Quinn, S. C. (2014). Early Intervention to Preempt Major Depression Among Older Black and White Adults. *Psychiatric Services, 65*, 765-773. doi:10.1176/appi.ps.201300216

Rizvi, S. L., Dimeff, L. A., Skutch, J., Carroll, D., & Linehan, M. M. (2011). A Pilot Study of the DBT Coach: An Interactive Mobile Phone Application for Individuals With Borderline Personality Disorder and Substance Use Disorder. *Behavior Therapy, 42*, 589-600. doi:10.1016/j.beth.2011.01.003

Robertson, R., Wenzel, L., Thompson, J., & Charles, A. (2017). *Understanding NHS financial pressures.* Technical report, The Kings Fund.

Rodda, J., Walker, Z., & Carter, J. (2011). CLINICAL REVIEW Depression in older adults. *BMJ, 343*. doi:10.1136/bmj.d5219

Rodriguez, J. D., Perez, A., & Lozano, J. A. (2010). Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*, 569-575. doi:10.1109/TPAMI.2009.187

Roy, T., & Lloyd, C. E. (2012). Epidemiology of depression and diabetes: a systematic review. *Journal of Affective Disorders, 142 Suppl*, S8--21. doi:10.1016/S0165-0327(12)70004-6

Rude, S. S., Durham-Fowler, J. A., Baum, E. S., Rooney, S. B., & Maestas, K. L. (2010). Self-report and cognitive processing measures of depressive. *34*, 107-115. S. S.

Sauve, L., Renaud, L., Kaufman, D., & Duplaa, E. (2016). Can Digital Games Help Seniors Improve their Quality of Life? *Proceedings of the 8th International Conference on Computer Supported Education, Vol 1 (Csedu)*, 94-102. doi:10.5220/0005800200940102

Scanlon, L., O'Shea, E., O'Caoimh, R., & Timmons, S. (2015). Technology use and frequency and self-rated skills: A survey of community-dwelling older adults. *Journal of the American Geriatric Society, 63*, 1483-1484.

Scazufca, M., P. de Paula Couto, M. C., Huang, H., Kester, R., Braga, P. E., Peluso, . . . Ribeiro, E. (2016). Public Stigma towards Older Adults with Depression: Findings from the São Paulo-Manaus Elderly in Primary Care Study. *PLOS ONE, 11*, e0157719. doi:10.1371/journal.pone.0157719

Schalinski, I., Teicher, M. H., Nischk, D., Hinderer, E., Müller, O., & Rockstroh, B. (2016). Type and timing of adverse childhood experiences differentially affect severity of PTSD, dissociative and depressive symptoms in adult inpatients. *BMC Psychiatry, 16*, 295. doi:10.1186/s12888-016-1004-5

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric Depression Scale: Recent evidence and development of a shorter version. In T.L. Brink (ed.) *Clinical Gerontology: A guide to assessment and intervention* (pp. 165-173). New York: The Haworth Press.

Sirey, J. A., Bruce, M. L., & Alexopoulos, G. S. (2005). The Treatment Initiation Program: an intervention to improve depression outcomes in older adults. *The American Journal of Psychiatry, 162*, 184-6. doi:10.1176/appi.ajp.162.1.184

Slade, M., Thornicroft, G., & Glover, G. (1999). The feasibility of routine outcome measures in mental health. *Social Psychiatry and Psychiatric Epidemiology, 34*, 243-249. doi:10.1007/s001270050139

Smyer, M. A., & Qualls, S. H. (1999). *Aging and Mental Health.* Oxford: Blackwell.

So, M., Yamaguchi, S., Hashimoto, S., Sado, M., Furukawa, T. A., & McCrone, P. (2013). Is computerised CBT really helpful for adult depression? - A meta-analytic re-evaluation of CCBT for adult depression in terms of clinical implementation and methodological validity. *BMC Psychiatry, 13*, 113. doi:10.1186/1471-244X-13-113

South Yorkshire and Bassetlaw Clinical Commissioning Group. (n.d.). Sustainability and Transformation Plan. Technical report.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder – the GAD7. *Archives of Internal Medicine, 166*, 1092-7. doi:10.1001/archinte.166.10.1092

Stanley, J. T., & Isaacowitz, D. M. (2011). Age-Related Differences in Profiles of Mood-Change Trajectories. *Developmental Psychology, 47*, 318-330.

The Academy of Medical Sciences. (2016). *Improving the Health of the Public by 2040: Optimising the research environment for a healthier, fairer future.* Technical report

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological), 58*, 267-288.

Tortajada, S., García-Gomez, J. M., Vicente, J., Sanjuán, J., Frutos, R., Martín-Santos, R., . . . Robles, M. (2009). Prediction of postpartum depression using multilayer perceptrons and pruning. *Methods of Information in Medicine, 48*, 291-8. doi:10.3414/ME0562

Unützer, J., Schoenbaum, M., Katon, W. J., Fan, M.-Y., Pincus, H. A., Hogan, D., & Taylor, J. (2009). Healthcare costs associated with depression in medically Ill fee-for-service medicare participants. *Journal of the American Geriatrics Society, 57*, 506-10. doi:10.1111/j.1532-5415.2008.02134.x

Vaportzis, E., Clausen, M. G., & Gow, A. J. (2017). Older Adults Perceptions of Technology and Barriers to Interacting with Tablet Computers: A Focus Group Study. *Frontiers in Psychology, 8*, 1687. doi:10.3389/fpsyg.2017.01687

Vis, C., Kleiboer, A., Prior, R., Bònes, E., Cavallo, M., Clark, S. A., . . . Riper, H. (2015). Implementing and up-scaling evidence-based eMental health in Europe: The study protocol for the MasterMind project. *Internet Interventions, 2*, 399-409. doi:10.1016/J.INVENT.2015.10.002

Vorrink, S. N., Antonietti, A. M., Kort, H. S., Troosters, T., Zanen, P., & Lammers, J.-W. J. (2016). Technology use by older adults in the Netherlands and its associations with demographics and health outcomes. *Assistive Technology*, 1-8. doi:10.1080/10400435.2016.1219885

Wardenaar, K. J., Loo, H. M., Cai, T., Fava, M., Gruber, M. J., Li, J., . . . Kessler, R. C. (2014). The effects of co-morbidity in defining major depression subtypes associated with long-term course and severity. *Psychological Medicine, 44*, 3289-3302. doi:http://dx.doi.org/10.1017/S0033291714000993

Watts, S., Mackenzie, A., Thomas, C., Griskaitis, A., Mewton, L., Williams, A., & Andrews, G. (2013). CBT for depression: a pilot RCT comparing mobile phone vs. computer. *BMC Psychiatry, 13*, 49. doi:10.1186/1471-244X-13-49

Webb, A. K., Jacobs-Lawson, J. M., & Waddell, E. L. (2009). Older adults' perceptions of mentally ill older adults. *Aging & Mental Health, 13*, 838-46. doi:10.1080/13607860903046586

Wernick, M., Yang, Y., Brankov, J., Yourganov, G., & Strother, S. (2010). Machine Learning in Medical Imaging. *IEEE Signal Processing Magazine, 27*, 25-38. doi:10.1109/MSP.2010.936730

Windle, K. (2015). What role can local and national supportive services play in supporting independent and healthy living in individuals 65 and over? Technical Report.

Yesavage, J. A., & Sheikh, J. I. (1983). Geriatric Depression Scale (GDS). *Clinical Gerontologist, 5*, 165-173. doi:10.1300/J018v05n01\_09

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of psychiatric research, 17*, 37-49.

Yochim, B. P., Mueller, A. E., & Segal, D. L. (2013). Late life anxiety is associated with decreased memory and executive functioning in community dwelling older adults. *Journal of Anxiety Disorders, 27*, 567-75. doi:10.1016/j.janxdis.2012.10.010

Yohannes, A. M., Baldwin, R. C., & Connolly, M. J. (2000). Depression and anxiety in elderly outpatients with chronic obstructive pulmonary disease: prevalence, and validation of the BASDEC screening questionnaire. *International journal of geriatric psychiatry, 15*, 1090-6.

Zbozinek, T. D., Rose, R. D., Wolitzky-Taylor, K. B., Sherbourne, C., Sullivan, G., Stein, M. B., . . . Craske, M. G. (2012). Diagnostic overlap of generalized anxiety disorder and major depressive disorder in a primary care sample. *Depression and Anxiety, 29*, 1065-71. doi:10.1002/da.22026

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica, 67*, 361-70.

# Appendices

## Appendix 1. Search terms for the first mapping review (Chapter 3A)

**WEB OF SCIENCE**

Date of search: 11/07/17

Topic and title searches. Under ‘more settings’, search language was set to English. Timespan: all years.

TS=(Machine learning or neural network$ or random forest$ or lasso)

AND

TS=(“Early sign$” or “early symptom$” or prodrome or “first symptom$” or predict or prediction)

AND

TS=(Depression or “depressive disorder$” or “depressive episode$” or anxiety or “generali$ed anxiety disorder”)

**EMBASE** via ovid

Date of search: 11/07/2017

Restricted to English language and EMBASE using search filters.

Machine learning or neural network$ or random forest$ or lasso

Or exp machine learning/

AND

Early sign$ or early symptom$ or prodrome or first symptom$ or predict or prediction

Or exp prodromal symptom/

AND

Depression or depressive disorder$ or depressive episode$ or anxiety or generali#ed anxiety disorder

Or exp depression/

**MEDLINE via Ovid**

Date of search: 11/07/17

(machine learning or neural network$ or random forest$ or lasso).mp.

or exp Machine Learning/ or exp "neural networks (computer)"/

AND

Early sign$ or early symptom$ or prodrome or first symptom$ or predict or prediction

AND

Depression or depressive disorder$ or depressive episode$ or anxiety or generali#ed anxiety disorder

Or exp Depression/ or exp Depressive Disorder/ or \*Anxiety/

Search limited to ‘humans’ and ‘English language only’.

**PSYCINFO**

Date of search 11/07/17

Machine learning or neural network$ or random forest$ or lasso

Or machine learning/ or artificial neural networks/ or learning/

AND

Early sign$ or early symptom$ or prodrome or first symptom$ or predict or prediction

Or prodrome/ or “onset (disorders)”/

AND

Depression or depressive disorder$ or depressive episode$ or anxiety or generali#ed anxiety disorder

Or exp major depression/ or exp anxiety/

Search limited to ‘Human’, ‘English language’, ‘All years’.

**ASSIA**

Date of search: 11/07/2017

“Machine learning” or “neural network\*” or “random forest\*” or lasso

(set to abstract)

AND

“Early sign\*” or “early symptom\*” or prodrome or “first symptom\*” or predict or prediction (set to abstract)

AND

Depression or “depressive disorder\*” or “depressive episode\*” or anxiety or “generali?ed anxiety disorder”

(set to abstract)

Limited to peer reviewed, all publication dates, English only, scholarly journals only.

## Appendix 2. Search terms for the second mapping review (Chapter 3B)

**PSYCINFO**

Date of search: 10/07/2017

Exp geriatrics/ or (Over 50? or older adult? or older people or older person? or senior? or silver surfer? or people aged over 50 or elderly or elder? Or geriatric).ab.

AND

(mobile phone application\*1 or cellphone application\*1 or cellular phone application\*1 or ipad application\*1 or tablet computer application\*1 or ipad app\*1 or tablet computer app\*1 or mobile phone app\*1 or smartphone app\*1 or smartphone\*1 or cellphone app\*1 or pc software or computer software or software or website\*1 or online or app\*1).ab.

AND

(perception? or experience? or attitude? or opinion?).ab.

AND

Exp mental health/ or exp affective disorders/ or exp anxiety disorders/ or exp anxiety/ or exp major depression/ or (depression or anxiety).ab.

Search limited to English language, human, 2007 to 2017.

**ASSIA**

Date of search: 10/07/2017

SU.EXACT("Older people") OR SU.EXACT(“Elderly People”) OR SU.EXACT(“Geriatrics”) OR “Over 50[\*1]” or “older adult[\*1]” or “older people” or “older person[\*1]” or senior[\*1] or “people aged over 50” or “silver surfer[\*1]”

AND

“mobile phone application[\*1]” or “cellphone application[\*1]” or “cellular phone application[\*1]” or “ipad application[\*1]” or “tablet computer application[\*1]” or “ipad app[\*1]” or “tablet computer app[\*1]” or “mobile phone app[\*1]” or “smartphone app[\*1]” or “smartphone[\*1]” or “cellphone app[\*1]” or “pc software” or “computer software” or “website[\*1]” or online or app[\*1] or software

AND

SU.EXACT("Mental health") OR SU.EXACT.EXPLODE("Depression") OR SU.EXACT.EXPLODE("Anxiety") OR SU.EXACT.EXPLODE("Affective disorders") OR “mental health” or depression or anxiety or “mental wellbeing” or “affective disorder[\*1]”

AND

Perception[\*1] OR experience[\*1] OR attitude[\*1] OR opinion[\*1]

Limited to peer-reviewed, limited to scholarly journals, limited to English language, past ten years.

Search in abstract only, in years 2007-2017

**SOCIOLOGICAL ABSTRACTS**

Date of search: 10/07/2017

SU.EXACT("Elderly") OR SU.EXACT(“Geriatrics”) OR SU.EXACT(“Gerontology”) “Over 50[\*1]” or “older adult[\*1]” or “older people” or “older person[\*1]” or senior[\*1] or “people aged over 50” or “silver surfer[\*1]”

AND

“mobile phone application[\*1]” or “cellphone application[\*1]” or “cellular phone application[\*1]” or “ipad application[\*1]” or “tablet computer application[\*1]” or “ipad app[\*1]” or “tablet computer app[\*1]” or “mobile phone app[\*1]” or “smartphone app[\*1]” or “smartphone[\*1]” or “cellphone app[\*1]” or “pc software” or “computer software” or “website[\*1]” or online or app[\*1] or software

AND

SU.EXACT("Mental health") OR SU.EXACT.EXPLODE("Depression (Psychology)") OR SU.EXACT.EXPLODE("Anxiety") OR SU.EXACT.EXPLODE("Affective Illness") OR “mental health” or depression or anxiety or “mental wellbeing” or “affective disorder[\*1]”

AND

Perception[\*1] OR experience[\*1] OR attitude[\*1] OR opinion[\*1]

Limited to peer-reviewed, limited to scholarly journals, limited to English language, past ten years. Search in abstract only.

**WEB OF SCIENCE (core collection)**

10/07/2017

TS=(“Over 50” or “over 50s” or “older adult$” or “older people” or “older person$” or “senior$” or “people aged over 50” or "elderly" or "geriatric")

AND

TS=(“mobile phone application$” or “cellphone application$” or “cellular phone application$” or “ipad application$” or “tablet computer application$” or “ipad app$” or “tablet computer app$” or “mobile phone app$” or “smartphone app$” or “smartphone app$” or “cellphone app$” or “pc software” or “computer software” or “website$” or “online” or app$ or software)

AND

TS=("Mental health" OR "Depression" OR "Anxiety" OR "Affective disorder$" OR “mental health” OR “mental wellbeing”)

AND

TS=(Perception$ OR experience$ OR attitude$ OR opinion$)

Timespan from 2007 to 2017, search language set to English.

217 results, imported to mendeley.

**MEDLINE**

Date of search: 10/07/2017

Exp geriatrics/ or (Over 50? or older adult? or older people or older person? or senior? or silver surfer? or people aged over 50 or elderly or elder? Or geriatric).ab

AND

Exp mobile applications/ or (mobile phone application\*1 or cellphone application\*1 or cellular phone application\*1 or ipad application\*1 or tablet computer application\*1 or ipad app\*1 or tablet computer app\*1 or mobile phone app\*1 or smartphone app\*1 or smartphone\*1 or cellphone app\*1 or pc software or computer software or website\*1 or online or app\*1 or software).ab

AND

(perception? or experience? or attitude? or opinion?).ab

AND

Exp mental health/ or exp mood disorders/ or exp anxiety disorders/ or \*anxiety/ or exp depressive disorder/ or exp depressive disorder, major/ or exp depression/ or (depression or anxiety).ab

Limited to English language, human, 2007 to 2017.

## Appendix 3. Activity sorting cards used in Study 2 (Chapter 5)

Cards were each printed with one of the following phrases, in size 55 text.

* Research a health problem
* Speak with a doctor
* Make a purchase with a credit card
* Read a self-help book
* Send a photo
* Keep a personal diary
* Relax and unwind
* Find a solution to a DIY problem
* Chat with friends
* Record your exercise

## Appendix 4. Vignette cards used in Study 2 (Chapter 5)

|  |  |
| --- | --- |
| A friend of yours has been feeling out of sorts recently. S/he has had a loss of appetite and complains of feeling tired a lot, although s/he finds it difficult to sleep at night. You also notice that s/he has been moving a lot slower than normal. S/he says that everything looks grey and washed out. | A neighbour of yours has been worrying a lot recently, making comments about the number of criminals in the area (despite living in a relatively safe area), talking a lot about it being flu season and also complaining that their cat seems unwell, even though the vet has said he’s fine. This neighbour seems quite agitated, and complains about insomnia. S/he has been quite irritable, which is not normal for them. |
| You wake up late in the morning and don’t feel like getting out of bed. When you finally do get up, you can find no pleasure in eating your breakfast, or going for a walk, which you used to enjoy. You had planned to play Bridge with some friends in the afternoon but you don’t go in the end because you don’t feel like you have the energy. There is a sense that a dark cloud is hanging over you. You’ve felt like this for some weeks now. | You get up in the morning and feel uneasy but you are not sure why. You notice it is quite cold and you are worried about getting ill. You go down stairs to put the heating on and you think you hear the boiler make a different noise to normal. You worry that something might be wrong with it. You set out your breakfast but you don’t feel hungry. Just sitting down makes you feel uncomfortable – you are very restless and yet you feel quite tired. You sense some pain in your legs. You have been having these kinds of experiences for over two weeks now. |

## Appendix 5. ScHARR Research Ethics Committee approval letter for Study 2

/Users/jakeandrews/Google Drive/Writing/Thesis/Approval Letter ethics 1.pdf

## Appendix 6. ScHARR Research Ethics Committee approval letter for Study 3

/Users/jakeandrews/Google Drive/Writing/Thesis/Approval Letter ethics 2.pdf

## Appendix 7. Health Research Authority approval letter for Study 5 (Chapter 8)

/Users/jakeandrews/Google Drive/Writing/Thesis/HRA approval page 1.pdf

## Appendix 8. Interview Schedule for Study 5 (Chapter 8)

1. Can you tell me a bit about your role and the type of patients you work with? (5-10 mins)

(*Follow up questions could ask about the participants’ experience with participants with mental health difficulties.)*

1. Do you have any experience with using technology/telemedecine with older patients? (1-10 mins)

*(Follow up questions could ask about how successful or unsuccessful such approaches were.)*

1. I am interested in the early detection of depression and anxiety in older people. I am developing software where older people report on their mood in their own homes once a day, every day, using a computer/tablet/mobile phone. The software then predicts whether the user may experience low mood or anxious feelings based on their responses. If it predicts they will, the software can provide an alert, which could be used for example to prompt a full screening to be undertaken. What are your immediate reactions to this? (5-10 mins)

*(Follow up questions could ask about more specific aspects of the approach, or ask participants to comment further based on their experience.)*

1. Thinking about the patient pathways you work in, could an alert system for early signs of mental health problems be easily added to these pathways to aid decision-making? (5 mins)

*(Follow up questions could ask about specific technologies or software currently used by participants that an alert system might be added to)*

1. Do you see any barriers to implementing this technology within the pathways you work in? (5-10 mins)

*(Follow up questions could ask about if participants have overseen the introduction of any other new technologies in the past and how these went)*