Identifying phenotypes and long term course of hand problems in older people using a latent transition approach

Daniel John Green

submitted for the degree of

Doctor of Philosophy

June 2016

Keele University

Declaration

This PhD project was funded by the National Institute for Health Research School for Primary Care Research (NIHR SPCR). The views expressed are those of the author and not necessarily those of the NHS, the NIHR or the Department of Health. This PhD project used data contained within the North Staffordshire Osteoarthritis Project (NorStOP), which was funded by a Medical Research Council (MRC) programme grant (grant code: G9900220), obtained by the Arthritis Research UK Primary Care Centre.

The initial PhD idea was developed by my original supervisors, Dr. Elaine Thomas and Prof. Danielle van der Windt, with input from myself and others within the Research Centre. However, throughout the course of the project, I have refined the focus of work, both clinically and methodologically, with guidance from all my supervisors over the 3 year period (Prof. Kelvin Jordan, Dr. Elaine Thomas, Prof. Danielle van der Windt and Dr. Jo Protheroe). The self-reported data in NorStOP was collected by the NorStOP project team, and the medical record information was collected by the informatics team at the Arthritis Research UK Primary Care Centre and the Keele GP Research Partnership. I have performed the investigations in this project using secondary analysis of previously collected data. I had no control over the content of the data collected.

Prof. Kelvin Jordan, Dr. Elaine Thomas, Prof. Danielle van der Windt and Dr. Jo Protheroe advised me on the planning of analyses, and the writing and presentation of chapters. All analyses (apart from the ALTA, clearly highlighted in *Chapter 9*) and written work has been conducted by myself. Dr. Bethany Bray (Penn State University), Dr. Brian Flaherty (Washington State University) provided advice via email or in person on various aspects of the LTA/ ALTA approaches. Dr. Linda Muthén and Prof. Bengt Muthén provided advice by email on the use of Mplus software.

Abstract

Musculoskeletal conditions of the hand are frequent causes of pain and disability in older people, yet knowledge regarding the characteristics and patterns of hand pain and problems over time is lacking. The objectives of this project were to identify sub-groups of older individuals with distinct presentations (phenotypes) of hand pain and function, investigate how these phenotypes changed over a 6 year period, and explore what characteristics and factors were associated with long-term outcomes. In addition to this, an exploration of the longitudinal association between hand phenotypes and mental health was performed.

The study population stemmed from the North Staffordshire Osteoarthritis Project (NorStOP); a large, general population-based, prospective cohort study of adults aged 50 years and over. Information on hand pain and problems was collected using questionnaires at baseline, 3 and 6 years. A total of 5,617 participants responded at all time points and were included in the analysis. Five phenotypes were identified using Latent Transition Analysis ('least affected', 'high pain', 'poor gross function', 'high pain and poor gross function' and 'severely affected') based on eight hand pain and function items. The most common transition between phenotypes was from 'high pain' at baseline to 'least affected' at 3 years. Individuals classified in the 'least affected' or 'severely affected' groups at baseline were the most stable. Individuals with nodes, chronic hand pain, sleep problems and bilateral hand pain at baseline were more likely to be in a more severe hand phenotype at 6 years. In addition to this, those in more severe hand phenotypes were more likely to have poor symptoms of mental health. The results provide clinically relevant information regarding the pattern of hand pain and problems over time, and the characteristics of those more likely to have an unfavourable outcome over a 6 year period.

Contents

List of Tables

List of Figures

Glossary

Abbreviations used within this thesis:

ACR= American College of Rheumatology

AIC= Akaike Information Criteria

AIMS2= Arthritis Impact Measurement Scale

ALTA= Associative Latent Transition Analysis

ANOVA= Analysis of Variance

APP= Average Posterior Probabilities

Anx/ dep= Anxiety and depression symptoms (mental health state)

AUSCAN= Australian/ Canadian hand osteoarthritis index

BAC= Blood Alcohol Content

BIC= Bayesian Information Criteria

BNF= British National Formulary

BM= Base Model

BMI= Body Mass Index

BMIs= Brief Motivational Interviews

CASHA= Clinical Assessment Study of the Hand

GLM= Generalised Linear Model

GMM= Growth Mixture Model

GP= General Practitioner

HADS= Hospital Anxiety and Depression Scale

High anx= High anxiety symptoms (mental health state)

HP= High Pain (hand state)

HPPGF= High Pain and Poor Gross Function (hand state)

HS= Health Survey

LA= Least Affected (hand state)

LCA= Latent Class Analysis

LCGA= Latent Class Growth Analysis

LLCA= Longitudinal Latent Class Analysis

LM= Latent Markov model

LPTA= Latent Profile Transition Analysis

LRT= Likelihood Ratio Test

LTA= Latent Transition Analysis

LTRA= Latent Transition Regression Analysis

MAR= Missing At Random

MCAR= Missing Completely At Random

Mild dep= Mild depression symptoms (mental health state)

MNAR= Missing Not At Random

MRI= Magnetic Resonance Imaging

MRM= Mixture Rasch Model

MSK= Musculoskeletal

NHS= National Health Service

NICE= The National Institute for Health and Care Excellence

No anx/ dep= No anxiety or depression symptoms (mental health state)

NorStOP= North Staffordshire Osteoarthritis Project

NSAIDs= Non-Steroidal Anti-Inflammatory Drugs

OA= Osteoarthritis

OR= Odds Ratio

PBIs= Parent-Based Interviews

PGF= Poor Gross Function (hand state)

RA= Rheumatoid Arthritis

RPS= Regional Pain Survey

RRR= Relative Risk Ratio

RUG= Research Users Group

SA= Severely Affected (hand state)

SAS= Statistical Analysis Software

SD= Standard Deviation

SEM= Structural Equation Modelling

SRM= Standard Response Mean

T1= Time 1 (and so on.)

UK= United Kingdom

WP= Widespread Pain

Acknowledgements

I would like to express my gratitude to my supervisors over the course of my PhD project Professor. Kelvin Jordan, Dr. Elaine Thomas, Professor. Danielle van der Windt and Dr. Jo Protheroe. Your skills, knowledge, patience, support, feedback and mentorship have been paramount throughout the PhD process. I would also like to thank Elaine and Danielle for selecting me for the position at Keele prior to the PhD. Many thanks to the members of the Research Users Group here at Keele, and Ethel Fish, for their time and input into my work.

I have immense gratitude for the help and guidance from Dr. Bethany Bray at Penn State University, USA, and Dr. Brian Flaherty at the University of Washington, Seattle, USA who on numerous occasions pointed me in the right direction, or provided clarification on the statistical analysis. Many thanks to Dr. Michelle Marshall for her advice and input on my project. Finally, many thanks to the NIHR NSPCR for selecting to fund this research.

Many thanks to the staff and students at Primary Care Sciences, Keele University for sharing their knowledge and kindness. Especially April, Di, Johny, Shula, Tom K and Tom P in the student room for their comradery, and the staff (although too many to mention), but especially Annette, Becky, Elaine N, Helen, Martyn, Trishna and Zoe for your companionship. Thanks to Mark Fudge for his support throughout my time at Keele.

I would also like to thank my Dad, Mum, siblings and family members for attempting to understand my project, and a huge thank you to all who proofread any of my work. Thanks to my nieces and nephews for maintaining my inner child and brightening my days. Thank you Alex, Kat, Katie, Luke and all at Hardwick Lane Farm for their warmth, help and friendship. Thanks to Craig, Christine and all at Dinnington RUFC for their determination and focus in my earlier years. And huge thanks to my partner Lou for always being there when I need you. Finally, thank you Nan for giving me the direction and drive to pursue.

“You can only do your best”

Edith Smith, my Nan.

“Progress is a game of inches, and leaps and bounds are only in hindsight.

[…] Sometimes science is literally like watching paint dry”

T Matalas/ T Fickett; 12 Monkeys

Publications/ Presentations arising from this thesis

Publication

Green DJ, Jordan KP, Protheroe J, van der Windt D. Development of hand phenotypes and changes in hand pain and problems over time in older people. PAIN 2016;157:569-76.

Oral Presentations

Green DJ, Thomas E, Protheroe J, van der Windt D. Statistical Methodology: A fresh approach to musculoskeletal research. May 2013: Post-Graduate Symposium, Primary Care Sciences, Keele University.

Green DJ, Thomas E, Protheroe J, van der Windt D. Latent Transition Analysis: what it is, and how I plan to use it. July 2013: Young Statistician's Meeting (YSM), Imperial College London.

Green DJ, Jordan KP, Protheroe J, van der Windt D. Defining latent phenotypes of patients with hand Osteoarthritis; a fresh approach to understanding musculoskeletal conditions. May 2014: Post-Graduate Symposium, Primary Care Sciences, Keele University.

Green DJ, Thomas E, Jordan KP, Protheroe J, van der Windt D. Defining latent phenotypes of patients with hand osteoarthritis (OA); a fresh approach to understanding musculoskeletal conditions. July 2014: Society for Academic Primary Care (SAPC), University of Edinburgh.

Green DJ, Jordan KP, Protheroe J, van der Windt D. "Sometimes, it's not 'one glove fits all'". Elevator pitch competition (awarded 2nd place). September 2014: NIHR/MRC Symposium for early career Research Methodologists, London.

Green DJ, Jordan KP, Protheroe J, van der Windt D. Developing osteoarthritis phenotypes and predicting changes in hand conditions over time in older people. May 2015: Post-Graduate Symposium, Primary Care Sciences, Keele University.

Green DJ, Jordan KP, Protheroe K, van der Windt D. The association between hand problems and mental health characteristics in older people over 3 years of follow up. October 2015: NAPCRG, Cancun, Mexico.

Poster presentations

Green DJ, Jordan KP, Protheroe J, van der Windt D. Modelling the long-term course and outcomes of hand pain and disability in older people with hand osteoarthritis (OA). September 2013: SPCR Trainee Meeting, University of Oxford.

Green DJ, Jordan KP, Protheroe J, van der Windt D. Modelling the long-term course and outcomes of hand pain and disability in older people with hand osteoarthritis (OA). September 2014: SPCR Trainee Meeting, University of Oxford.

Green DJ, Jordan KP, Protheroe J, van der Windt D. Modelling the long-term course and outcomes of hand pain and disability in older people with hand osteoarthritis (OA). September 2014: SPCR Showcase, University of Oxford.

Green DJ, Jordan KP, Protheroe J, van der Windt D. Developing osteoarthritis (OA) phenotypes and predicting changes in hand conditions over time in older people. OARSI May 2015: Seattle, USA.

Green DJ, Jordan KP, Protheroe J, van der Windt D. Developing osteoarthritis (OA) phenotypes and predicting changes in hand conditions over time in older people. May 2015: Post-Graduate Symposium, Primary Care Sciences, Keele University.

Chapter 1: **Introduction**

This study examines the long term course of hand pain and hand function problems in older people, applying a longitudinal statistical method called Latent Transition Analysis.

1.1 Hand problems in older people

Pain, stiffness, and functional difficulties related to the hand are common in older people and can be due to a range of conditions, with osteoarthritis (OA) being the most frequent cause of pain and disability (Dziedzic et al., 2007; Spies-Dorgelo et al., 2007; Marshall, 2010; Marshall et al., 2013; Kloppenburg et al., 2014). In the older population (aged 50 years and older) the one year period prevalence of hand pain due to musculoskeletal (MSK) conditions has been estimated at 31%-47%, hand function difficulties at 47%-59% and hand stiffness at 39% (Dziedzic et al., 2007; Bellamy et al., 2011). Functional difficulties with the hand can have a severe impact on the everyday life of individuals, can limit participation in occupational and social activities, and be related to increased mental health problems (Kjeken et al., 2005; Hill et al., 2011; Bukhave and Huniche, 2014). Typical everyday issues with function include activities that require substantial grip strength coupled with twisting, difficulties getting dressed and writing, and handling or fingering small objects (Zhang et al., 2002; Kjeken et al., 2005). Not all hand conditions and experiences are the same (Kloppenburg and Kwok, 2011), and people experience different levels of hand joint pain, stiffness and functional difficulties, whilst bilateral (both hands) symptoms may occur (Zhang et al., 2002).

There is limited knowledge regarding the course of hand conditions over long-term follow-up, the different profiles of hand pain/ problems and the factors that are predictive of future

outcomes (Kloppenburg et al., 2007; Nicholls et al., 2012). In addition to this, while research into hand OA/ pain/ problems has increased over the last ten years, studies have highlighted a lack of understanding of the disease condition and relevant outcome measures (Kloppenburg et al., 2007; Dziedzic, 2013; Kloppenburg et al., 2014). Results from prognosis research, such as investigating the likely course of pain and function, and investigating the predictive value of key prognostic factors of interest, can provide more detailed information for researchers and clinicians.

Increasing knowledge about prognosis can help researchers to tailor future research objectives and inform study design by providing information on the expected frequency of outcome in cohorts and trials, therefore supporting the identification of individuals more at risk of future deterioration, and those most at need of intervention. Increasing the knowledge for clinicians, who are frequently the direct point of contact for patients, may support their decision making regarding the need for treatments or referral in patients with hand pain/ problems (i.e. which patients are most likely to benefit from intervention). For example, the research may reveal individuals most at risk of poor long-term outcomes of pain and disability, who can therefore be targeted with more treatment and monitoring. Ultimately, identifying the types of hand pain/ problems that have a worse prognosis could be crucial for tailoring treatment and referral options to those most affected by the condition.

Previous research on profiles of hand pain/ problems is limited, with the majority of studies investigating pain and function as separate entities (Jones et al., 2001; Zhang et al., 2002; Dominick et al., 2005; Kjeken et al., 2005; Dillon et al., 2007; Grotle et al., 2008a; Spies-Dorgelo et al., 2008; Botha-Scheepers et al., 2009; Bijsterbosch et al., 2011; Ghosh et al., 2014), while some studies have analysed total AUSCAN (Australian/ Canadian Hand Osteoarthritis Index, described in *section 4.3.3*) questionnaire scores to reflect an

overall score for pain and function combined (Allen et al., 2006a; Kim et al., 2010; Haugen et al., 2011b). Zhang and colleagues investigated more detailed aspects of individual function limitations in older people with symptomatic hand OA but pain was not explicitly investigated (Zhang et al., 2002). Kjeken and colleagues used scores of pain during activities as an unadjusted predictor of AUSCAN function, finding a positive relationship (more pain leads to more hand dysfunction), however detailed function aspects were the core focus of that study (Kjeken et al., 2005). It is a reasonable assumption that pain and function may be directly related to each other, and both aspects are important to individuals with hand OA, therefore both of these elements should be included in research regarding the longitudinal patterns of hand problems.

Further to this, longitudinal patterns of pain and function scores have predominately been analysed using the mean change between two time points (Botha-Scheepers et al., 2009; Bijsterbosch et al., 2011), or used as factors in regression models (using pain or function scores to predict presence of hand pain/ problems/ OA, Grotle et al., 2008a; Spies-Dorgelo et al., 2008; Kim et al., 2010). These approaches cannot fully explain the potentially complex relationships between pain and function and how this relationship might change over time. Evidently there is a significant gap in the literature regarding the profiles of both hand pain and hand function aspects in older people, and the way these have been investigated over time, which this project will address.

1.2 Latent Transition Analysis

As hand pain/ problems in older adults are unlikely to be homogenous, and individuals may report different profiles of pain and function severity, a more adaptive technique is needed that can move research forward from assessing changes in mean scores of pain and function independently. In addition, a technique that has the ability to model movement of

individuals between profiles over time would be useful to increase the knowledge of the likely course of hand pain and function in individuals. One such approach is Latent Transition Analysis (LTA). LTA will be explained in detail in *Chapter 3*, however briefly, LTA is a longitudinal method that has the ability to define distinct profiles (states) based on characteristics of interest, such as hand pain and different measures of function, whilst also incorporating change over time (Collins and Wugalter, 1992; Collins and Lanza, 2010). This approach designates each individual into one state (and only one state) at baseline and permits each individual to change their state membership at each follow-up time point investigated. By implementing this technique, it may be possible to identify distinct profiles of hand pain and function, and explore the likely pattern of change (transition) over time for each individual between the identified states. Therefore, a focus of this thesis will explore the application of LTA to investigate the long-term course of pain and function in older people with hand pain/ problems.

While this thesis has clear clinical objectives, the project will also explore methodological challenges to the statistical technique of LTA, by investigating relevant extensions commonly, and uncommonly, used by other researchers. The technique of LTA is a relatively novel approach to investigating change over time compared to more established statistical approaches (such as regression models and related latent class analysis), and as presented in later chapters certain aspects of the technique have not been widely explored. Therefore, this project will also address these alternative aspects of LTA, whilst reflecting on its suitability to investigate the longitudinal patterns of hand conditions and musculoskeletal conditions in general.

1.3 Research objectives

The intention of this section is to provide an overview of the research objectives of this PhD thesis, whilst providing a brief summary of the study used for the project. The overall aim of this thesis is to provide evidence on the longitudinal profiles of hand pain/ problems in older people, and the characteristics associated with more unfavourable outcomes. In order to achieve this aim, five main objectives have been developed, each of which has its own clinical questions with appropriate methodological approaches.

1.3.1 Data

The data that will be used for these objectives is explained in more detail in *Chapter 4, section 4.3* however, to provide a context, a brief summary is presented here.

All the objectives of the project will be examined using data from the NorStOP (North Staffordshire Osteoarthritis Project), which is a large observational cohort of the general population aged 50 years and over (Thomas et al., 2004). The participants were patients registered at one of eight general practices in North Staffordshire. Patients were sent a Health Survey (HS) questionnaire and if they indicated that they had experienced some pain or problems in their hand, hip, knee or foot over the last 12 months, the respondent was then sent a more detailed Regional Pain Survey (RPS). This questionnaire included more detailed questions related to each respective joint, and in the case of the hand it specifically included the AUSCAN (Australian/ Canadian Hand Osteoarthritis Index, Bellamy et al., 2002a) and the AIMS2 (Arthritis Impact Measurement Scales, Meenan et al., 1992). This process of a HS questionnaire, followed by a RPS questionnaire for those with recent joint problems was repeated at 3 years and 6 years follow-up.

1.3.2 Objective 1: Develop distinct states of hand pain and hand function

The first objective is to define distinct profiles (states) of hand pain and function in a community-based sample of older adults based on responses to hand pain and function statements measured at baseline, 3 years and 6 years. In this project, the statements used to identify the states are termed 'indicators' and this terminology will be used from this point forward.

Specific objectives are to:

- Identify hand pain and function issues relevant to individuals with hand pain/ problems;
- Identify common states of hand pain, function and stiffness based on responses to key indicators of interest;
- Assess the use of LTA as an approach to identifying states of hand pain and function.

To address this first objective, key indicators from the NorStOP questionnaires will be used and applied within LTA to identify distinct states of hand pain/ problems (*Chapter 4*). These developed states could potentially reflect individuals affected by combinations of pain (at rest or during activities), weak grip strength, poor fine motor skills, restriction on daily activities or various severities of stiffness. The robustness and goodness of fit of the model will also be tested (*Chapter 5*). The latent model and states that are developed will form the basis for the majority of the work contained in this PhD thesis.

1.3.3 Objective 2: Investigate the transitions of individuals between the hand states over time

The overall focus of this objective is to determine the likelihood of transitions between the hand states over the 6 years follow-up period. An assessment will be made to determine if probability of transitions vary between characteristics, such as gender or age.

Specific objectives are to:

- Determine the transition probabilities between hand states over the 6 years follow-up period;
- Investigate whether transition probabilities are significantly different for key demographic factors (age, gender, widespread pain, living status);
- Apply extensions to LTA, to stratify transition probability by demographic factors, and investigate whether the inclusion of covariates into the LTA modelling process is necessary to improve model fit.

1.3.4 Objective 3: Explore factors predicting future hand state membership

Specific objectives are to:

- Explore which sociodemographic, general health and hand-related factors are most likely to predict future state membership;
- Explore whether these factors are associated with improvement or deterioration of hand conditions over the 6 year period.

Baseline factors will be used to predict hand state membership at 3 and 6 years, therefore identifying characteristics for General Practitioners (GPs), other clinicians and researchers to be aware of in individuals with hand pain/ problems (*Chapter 7*). Particular hypotheses

of interest in this section are whether those with established risk factors for hand pain/ problems in older people, such as gender, age, obesity, excessive use of hand and longer symptom duration are associated with more severe state membership.

1.3.5 Objective 4: Investigate the associations of hand states with primary care consultation and medication prescription

Specific objectives are to:

- Explore whether hand state membership at baseline is associated with primary care consultation for hand-related conditions, and subsequent medication prescription;
- Explore if the type of medication prescribed is associated with future hand state.

Therefore, the main focus of this objective is to investigate the proportion of individuals consulting for hand-related conditions and then prescribed analgesia over the 6 years follow-up period, and explore whether this is associated with hand state severity (*Chapter 8*). The focus here is to assess whether individuals with hand pain/ problems are accessing health care, and to explore whether those receiving medications appear to benefit (in terms of symptom improvement/ deterioration). Should individuals with hand pain/ problems not consult their GP, this could potentially reflect a lack of belief that GPs can help from the patient perspective, or that patients feel they can manage their hand pain/ problem without medical help. Similarly, should those receiving the relevant prescriptions for hand pain/ problems not experience improvements, or rather continue deteriorating, this could highlight the need for further treatment development for hand conditions in older people.

1.3.6 Objective 5: Explore the longitudinal relationship between hand states and mental health states

Specific objectives are to:

- Assess the stability of mental health symptoms (anxiety, depression) in older people over the 6 years follow-up;
- Assess the use of ALTA (Associative Latent Transition Analysis) to explore the longitudinal relationship between hand states and mental health states over the 6 years follow-up period.

A strong relationship between mental health and musculoskeletal conditions, especially pain, has been found in previous research. However, little work has focussed on the association between mental health and hand pain/ problems in older people. The main focus of this objective is to develop latent states of anxiety and depression using LTA and use the novel extension of ALTA (further explained in *section 3.9.6*) to explore the longitudinal relationships between hand and mental health (*Chapter 9*).

Specific questions of interest for this objective are:

- Are those who are in more severe hand states more likely to be in more severe mental health states, and also more likely to remain in those severe states at follow-up?
- Are individuals who have improvements in their hand symptoms, also likely to have improvements in their mental health?

Exploring longitudinal relationships between hand and mental health will provide information on whether researchers and clinicians should be aware of mental health symptoms in older people with hand pain/ problems.

1.4 Outline of the thesis

The statistical technique of LTA and the results developed from it will be used throughout the thesis. **Chapter 2** will present an overview of the current literature of hand pain/

problems/ OA in older people, presenting information on epidemiology, diagnosis, risk factors and management. In **Chapter 3**, LTA is described followed by description of an applied example of the technique. After this, the results of a literature review to explore the current fields that LTA has been used in and application of relevant extensions to the method will be detailed. **Chapter 4** will describe the NorStOP study forming the basis for the analysis in this project and present the development of a base model to identify the states of hand pain/ problems in older people. The validity of this model will be assessed in **Chapter 5**, along with how hand state definition might potentially change over the 6 years follow-up period. Transitions between different states of hand pain and function will be presented in **Chapter 6**, with relevant extensions of the method to further understand the characteristics of individuals more/ less likely to transition.

In **Chapter 7**, the association of potential baseline factors, including demographic, general health and hand factors, with hand outcomes at the 3 and 6 years follow-up will be explored. The relationship of the identified hand states with health care use and pharmacological management will be assessed in **Chapter 8**. The longitudinal relationship between hand pain/ problems and mental health (symptoms of anxiety and depression) is assessed in **Chapter 9** where the application of a novel complex adaptation of LTA (Associative Latent Transition Analysis (ALTA)) is explored. Sensitivity analyses for the model developments in the thesis are reported in **Chapter 10**, before the thesis is summarised and concluded in **Chapter 11**.

Chapter 2: **Background: Hand problems in older people**

2.1 Objective

This chapter describes current knowledge on the epidemiology, diagnosis, risk factors and the long-term prognosis of hand conditions in older people. The main focus of this chapter is to provide an overview of the current understanding of hand pain/ problems in older people.

2.2 Epidemiology

The participants analysed in this project provided self-reported information on pain and functional difficulties in hand joints and are aged 50 years and older. Such hand problems in older individuals are most likely to be associated with a clinical diagnosis of hand OA (Spies-Dorgelo et al., 2007; Marshall, 2010; Marshall et al., 2013), and more broadly, joint pain in people over 45 is considered likely to be due to OA (NICE guidelines, 2014). Therefore, the information presented here includes information on individuals with hand OA, as well as those with hand pain and hand problems (pain/ problems) in general. The prevalence of hand OA varies greatly dependent on the definition of hand OA (whether confirmed by x-ray or meeting clinical criteria without an x-ray, described in *section 2.3*). Additionally, prevalence varies depending on the population used, and the ages of the participants (*Table 2.2.1*).

Table 2.2.1: Summary of the prevalence of clinical and radiographic OA as estimated in the literature.

Study/ Year	Country	Mean age (years)	Prev. of symptomatic HOA			Prev. of radiographic HOA		
			Overall	Female	Male	Overall	Female	Male
Carman et al., 1994	US	38				41%		
Chaisson et al., 1997	US	79-100				92%		
Caspi et al., 2001	Israel	79	77%	75%	81%	83%	83%	83%
Jones et al., 2001	Tasmania	56					56%	36%
Zhang et al., 2002	US	>70		26%	13%			
Aihie Sayer et al., 2003	UK	53		4%	1%		30%	19%
Haara et al., 2003	Finland	>30					48% (finger)	44% (finger)
Zhang et al., 2003	China	>60		6%	3%		47%	45%
Dahaghin et al., 2005	Nether-lands	67				62%	67%	55%
Rossignol et al., 2005	France	62	27%	43%	16%			
Dillon et al., 2007	US	>60	8%	9%	7%			
Spies-Dorgelo et al., 2007	Nether-lands	49	17%					
Grotle et al., 2008a	Norway	45	4.3%	5.8%	2.5%			
Botha-Scheepers et al., 2009	Nether-lands	61				75%		
Bernard et al., 2010	US	63				42%		
Bijsterbosch et al., 2011	Nether-lands	60				71% (nodal)		

Study/ Year	Country	Mean age (years)	Prev. of symptomatic HOA			Prev. of radiographic HOA		
			Overall	Female	Male	Overall	Female	Male
Haugen et al., 2011a	US	59					44%	38%
Massengale et al., 2012	US	70	8.5%	10%	7%			
Ghosh et al., 2014	India	>40				28%		
Prieto-Alhambra et al., 2014	Spain	64	2.4*	3.5*	1.3*			
Visser et al., 2014	Nether-lands	56	14%	20%	8%			
Yeşil et al., 2014	Turkey	54	3%	4%	0%			

Footnote: * prevalence per 1,000 person years; prev.= prevalence; HOA= Hand Osteoarthritis.

Table 2.2.1 displays a summary of reported prevalences for different nationalities, ages, gender and OA definition (the list of studies listed is not exhaustive). It can be seen that the prevalence for radiographic hand OA tends to be much higher than for clinical hand OA, and higher for older ages and female gender.

2.3 Diagnosis/ classification of hand OA

Various criteria have been proposed for the diagnosis of hand OA, based on results from clinical history, examination and/ or radiographs (x-ray/ MRI (Magnetic Resonance Imaging) scans) (Zhang et al., 2009). The most frequently used criteria for hand OA are based on the American College of Rheumatology (ACR) classification for symptomatic OA of the hand (Altman et al., 1990), which have been summarised in *Table 2.3.1* below. Hand OA is characterised by hand pain or stiffness, as well as swelling, enlargements or deformities of particular joints in the hand, which can be identified during a clinical examination (Altman et al., 1990; Kloppenburg and Kwok, 2011). The joints most often

affected are the: second and third distal interphalangeal (DIP) joints, which are the finger joints nearest the fingertips; second and third proximal interphalangeal (PIP) joints, which are the joints in the middle of the finger (the finger joint between the DIP and the metacarpophalangeal joint (MCPJ)/ 'knuckle'); and the first carpometacarpal (CMC) joint, which is the joint at the base of the thumb at the wrist (Altman et al., 1990; Kwok, 2013). Individuals diagnosed by the ACR criteria (i.e. without an x-ray) have been referred to in the literature as having symptomatic/ clinical hand OA (Felson and Nevitt, 2004; Kalichman and Hernández-Molina, 2010; Haugen et al., 2011a).

Table 2.3.1: American College of Rheumatology (ACR) classification criteria for osteoarthritis of the hand (Altman et al., 1990).

Hand pain, aching or stiffness AND three or four of the following features:
Hard tissue enlargement of two or more of 10 selected joints*
Hard tissue enlargement of two or more DIP joints
Fewer than three swollen MCP joints
Deformity of a least one of 10 selected joints*

Footnote: *: The 10 selected joints are the second and third distal interphalangeal (DIP), the second and third proximal interphalangeal (PIP), and the first carpometacarpal joints of both hands; DIP= distal interphalangeal; MCP= metacarpophalangeal.

Alternatively, individuals may receive a diagnosis of radiographic hand OA, which is most commonly assessed through the Kellgren-Lawrence (KL) grading scheme after an x-ray (Zhang et al., 2009). Generally, those with at least one joint with a KL grade ≥2 are considered to meet the definition of having radiographic hand OA (Zhang et al., 2009; Kalichman and Hernández-Molina, 2010; Haugen et al., 2011a), however some variations in the number of affected joints required, or inclusion of specific joints have been used in previous literature (Marshall et al., 2008). In addition to this, a EULAR (European League Against Rheumatism) working group has proposed a classification of hand OA based on physical findings and radiographic patterns (Zhang et al., 2009).

Radiographic hand OA can be divided into subsets of various profiles based on radiographic findings, such as erosive OA, thumb-based OA, nodal OA, non-nodal OA and generalised OA (Zhang et al., 2009; Kloppenburg and Kwok, 2011; Marshall et al., 2013). However, research into whether these subsets represent distinct entities or presentations of varying severity of the condition is still ongoing (Verbruggen et al., 1996; Punzi et al., 2010; Marshall et al., 2013).

Most clinical guidelines do not recommend taking radiographs as common practice as a clinical examination is often sufficient for diagnosis and they are not needed for deciding early/ core treatments for OA; this avoids potential side-effects of radiation (NICE (National Institute for Health and Care Excellence) guidelines, 2014). As described in *section 2.2*, the prevalence of symptomatic and radiographic hand OA can vary widely, with the prevalence of symptomatic hand OA generally ranging from 1% to 26% (Zhang et al., 2002; Aihie Sayer et al., 2003), while radiographic OA prevalence can range from 29% to 92% in the older population (Chaisson et al., 1997; Lawrence et al., 2008). A systematic review comparing associations between symptomatic and radiographic hand OA identified a positive association between radiographic hand OA and symptomatic hand pain (Dahaghin et al., 2006). However, inconsistent evidence was found for the relationship between radiographic hand OA and symptomatic hand function (Dahaghin et al., 2006). A further study found self-reported changes in pain and function scores were not associated with radiographic progression; however, this study was relatively short-term with only 2 years of follow-up (Botha-Scheepers et al., 2009).

Hand OA is the most common likely diagnosis for older individuals with hand pain and functional difficulties, however other diagnoses can present with pain and/ or function problems (Spies-Dorgelo et al., 2008; Marshall, 2010; Marshall et al., 2013). These include carpal tunnel syndrome (characterised by pain, tingling (pins and needles) and numbness in

the first three fingers, and thumb weakness), trigger finger (characterised by the locking of fingers/ thumbs when the digit is bent towards the hand), or DeQuervain tenosynovitis (characterised by pain and burning sensations in the thumb and wrist, with difficulty gripping). While these other conditions could be present in a small subset of participants contained in the study population, these were found to be uncommon compared to an OA diagnosis in a subsample of patients from NorStOP (Marshall et al., 2013).

Therefore, there are some discrepancies in diagnosis and classification of hand OA in the current literature, and in the resulting prevalence estimates. While assessing radiographic outcomes can be regarded as more objective than self-reported outcomes, it can be argued that the emphasis for hand OA (and OA in general) should be placed on patient reported levels of pain and function as the primary focus (Felson and Nevitt, 2004). Exploring the self-reported measures that have a direct impact on the everyday life of individuals (such as pain and function) should be prioritised, because improvements in these measures would provide more desirable benefits to affected individuals.

2.4 Risk factors for hand OA

The main risk factor for hand OA is age (Caspi et al., 2001; Haara et al., 2003; Grotle et al., 2008a; Ghosh et al., 2014; Prieto-Alhambra et al., 2014). Previous work has indicated that as many as 92% of older individuals have evidence of radiographic hand OA (Chaisson et al., 1997), and it has also been suggested elsewhere that the majority of older adults will have some evidence of radiographic deterioration (Felson and Nevitt, 2004). However, some studies have identified substantially lower rates of symptomatic hand OA classification (around one in five older female individuals, Visser et al., 2014). In addition, some studies have revealed that women aged between 50 to 65 years old (average post-menopausal years) are most at risk of developing hand OA compared to other years, and

men (Kalichman and Kobyliansky, 2007; Prieto-Alhambra et al., 2014; Yeşil et al., 2014). Therefore, this highlights that a more complex relationship between age and the development of hand OA may be present.

In addition to age, gender differences have been found in many studies; as mentioned above, the majority highlighting a higher prevalence of hand OA for women (van Saase et al., 1989; Zhang et al., 2002; Aihie Sayer et al., 2003; Zhang et al., 2003; Visser et al., 2014). A study of Russian women found that women with larger, more athletic builds and factors associated with increased rates of ageing process (such as younger age at menarche) were more likely to develop hand OA (Kalichman and Kobyliansky, 2007). However, while hand OA prevalence in women appears to peak at particular ages around 50 to 65 years, risk for men appears to constantly increase with age (Kallman et al, 1990; Prieto-Alhambra et al., 2014). Additional work in a UK cohort has found that males who were heavier at birth were slightly more likely to develop hand OA in later life, while a subset of men born lighter, but heaviest in the cohort at 53 years of age were more than twice as likely to develop hand OA (Aihie Sayer et al., 2003). In addition to this, Chaisson and colleagues reported from the Framingham study that men with higher maximal grip strength were at a higher risk of hand OA (Chaisson et al., 1999).

The relationship between obesity and hand OA is unclear with publications indicating obesity as a predictor of hand OA in many different populations (Carmen et al., 1994; Haara et al., 2003; Kalichman and Kobyliansky, 2007; Bernard et al., 2010; Ghosh et al., 2014; Prieto-Alhambra et al., 2014), while some have found no evidence of an association (Caspi et al., 2001; Marshall et al., 2013). Rather than using Body Mass Index (BMI), researchers have found associations between other measures of obesity, such as fat percentage, waist-to-hip ratio and waist circumference, with hand OA (Kalichman and Kobyliansky, 2007; Visser et al., 2014). Additional research has found metabolic factors,

such as serum glucose and insulin concentrations, to be associated with hand OA, indicating a potential metabolic relationship between obesity and hand OA (Visser et al., 2013). A further analysis of the NHANES III cohort found no association between hand OA and serum leptin (a hormone produced by fat cells) concentration levels (Massengale et al., 2012). However, another study which found a significant positive relationship between obesity and hand OA in Norway, also found a predictive (albeit not significant) association of those underweight (BMI <20) with the onset of OA (Grotle et al., 2008b). Therefore, while associations have been found, the relationship between hand OA and obesity is still to be fully understood.

Previous research investigating other risk factors has identified those with more manual occupations (similar to other findings related to grip strength and athletic build) to be more at risk of hand OA (Haara et al., 2003; Rossignol et al., 2005; Bernard et al., 2010). Lifestyle factors such as smoking and alcohol consumption have been found to have no association with presence of hand OA (Hart and Spector, 1993; Kalichman et al., 2005; Yeşil et al., 2014), while frequent chopstick use has been found to be significantly associated with the prevalence of hand OA in China (Hunter et al., 2004). Some evidence has highlighted a relationship between hand OA and knee OA, however this association is still not understood (Dieppe et al., 2000; Ghosh et al., 2014; Prieto-Alhambra et al., 2014). Other work investigating medical factors in relation to hand OA has found lower bone mineral density in the forearm in post-menopausal women with hand OA (Kim et al., 2010) and higher rates of carotid plaques and coronary calcification (potentially linked to metabolic factors previously mentioned) in older women with hand OA (Jonsson et al., 2009).

Much work has been completed on investigating the relationship between hand OA and specific genes, for example in the Framingham study, the GOGO study and the GARP

study (Spector et al., 1996; Felson et al., 1998; Demissie et al., 2002; Kalichman et al., 2003; Kraus et al., 2007; Bijsterbosch et al., 2013; Styrkarsdottir et al., 2014), and also the hereditary nature of hand OA (Riyazi et al., 2005; Ghosh et al., 2014). However, the role of genetic factors in the development of hand OA is still to be totally understood (Leung et al., 2013) and is beyond the scope of this PhD project.

2.5 Management options

The management of hand OA in older people consists of pharmacological (such as analgesia, non-steroidal anti-inflammatory drugs (NSAIDs), creams (topical NSAIDs)), and non-pharmacological approaches including heat/ cold therapy and advice on exercises (Zhang et al., 2007; NICE guidelines, 2014). For individuals consulting with hand OA in the UK, the recommended approach taken by the GP is to first discuss exercise options, along with paracetamol if the patient believes it is beneficial (NICE guidelines, 2014). Should these steps be unsuccessful in reducing pain and/ or improving function and daily activities, the prescription of NSAIDs and then opioid analgesia can be considered (NICE guidelines, 2014).

NSAIDs have been shown to result in some improvements in pain and function, however some concerns remain due to potential side effects such as gastrointestinal, renal, cardiovascular, liver and skin problems (Hungin et al., 2001; Kean et al., 2008; Koffeman et al., 2015). Further to this, the efficacy of strong opioid analgesia for musculoskeletal problems is questionable, and side-effects in the elderly population are more prevalent (Saunders et al., 2010; Ashworth et al., 2013).

Clinical trials have found improvements in pain and function after 12 months of using a night thumb splint (Rannou et al., 2009), while another trial provided evidence of an improvement in hand strength following hand exercises and hand protection (Stamm et al.,

2002). A more recent trial focussing on hand exercise found a modest improvement in hand pain and stiffness at 3 months, however these findings were not sustained through to 6 months (Østerås et al., 2014). Aside from this, a trial in older adults with hand OA evaluating support by occupational therapists in self-management and joint protection revealed some modest improvements in pain over a 12 month period (Dziedzic et al., 2013). However, the effectiveness of these interventions needs to be supported by further evidence and are yet to reach widespread general practice.

Qualitative research has identified that older UK individuals with hand OA believe that there are a lack of treatment options available, while some received contradictory advice about the management of the condition (Hill et al., 2011). However, further qualitative research revealed, potentially as a consequence, that individuals do not frequently seek professional/ clinical help for their problems, or do not regularly take the medication prescribed for their hand condition (Hill et al., 2011; Bukhave and Huniche, 2014). Other work has highlighted that many individuals adapt their lifestyle, through either the use of assistive devices, avoiding activities that could trigger pain or functional issues, or seeking assistance from others to carry out tasks (Myers 2008; Bukhave and Huniche, 2014). A clear effective treatment for older individuals with hand OA/ pain/ problems is not currently available, or confirmed (Kloppenburg et al., 2014). Current guidance in the UK, US and Europe is for exercise advice (which has been shown to have small short-term benefits), oral medication (Zhang et al., 2007; Hochberg et al., 2012; NICE guidelines, 2014), which is standard for all OA conditions (but with concerns surrounding the efficacy and safety of pharmacological treatments) and topical NSAIDs (creams), which has less safety risk compared to oral NSAIDs.

2.6 Long-term prognosis and prognostic factors

Studies investigating the long term course of hand pain/ problems report wide variation in the prognosis of the condition. In a study of adults (18 years and older) consulting with hand and wrist problems in general practice, 42% reported a complete recovery at 12 months, while 37% reported little or no recovery (Spies-Dorgelo et al., 2008). Factors that influenced a poor outcome were female gender, older age, longer symptom duration (>3 months) and lower coping strategies (Spies-Dorgelo et al., 2008), similar to factors highlighted in the systematic review by Nicholls et al., 2012. However, as also highlighted by Nicholls et al. individuals consulting for hand pain/ problems may reflect a population with more severe hand symptoms, and therefore studies based in the general population would capture a broader spectrum of hand symptom severities (Nicholls et al., 2012). In addition, Spies-Dorgelo et al. used a younger population, which therefore may represent individuals more amenable to improvements compared to the elderly (Spies-Dorgelo et al., 2008).

A 2 year study of radiological and clinical changes in hand OA found increases in self-reported pain, function, and pain intensity mean scores, as well as radiographic deterioration over the follow-up period (Botha-Scheepers et al., 2009). Those who displayed radiographic deterioration were more likely to be older, female and in the early years (<10) post-menopause. Interestingly, the authors found that changes in self-reported pain and function scores over the 2 years were not associated with radiographic progression. An extended study of the same cohort investigating radiographic change over a 6 year period in individuals with confirmed hand OA found that radiographic progression was present in 52.5%, and progression was associated with the baseline predictors of pain and presence of nodes (Bijsterbosch et al., 2011).

A UK cohort (the Bristol 'OA500 study') revealed that individuals with radiographic hand OA deteriorated in terms of self-reported pain and overall change over an 8 year period (Dieppe et al., 2000). In addition, the authors also found that of those with hand OA alone at baseline, 44% had acquired knee or hip OA 8 years later. Finally, in a 9-year (Framingham Osteoarthritis) study based on individuals with radiographic hand OA at baseline (mean age 59 years), 96% of women and 91% of men showed radiographic progression over the 9 year period (Haugen et al., 2011a).

The results from these longitudinal studies indicate that many individuals with hand OA, whether defined as symptomatic or radiographic OA, deteriorate over time. In particular, those with radiographic OA appear less likely to improve than those with symptomatic OA, but the characteristics of those who do improve are generally unclear. It would be of interest to identify the characteristics of those who improve, and deteriorate, over a long-term follow-up to provide further knowledge regarding the prognosis of symptomatic hand OA.

This chapter has presented an overview of the current literature of hand conditions/ OA in the older population. The next chapter will present a detailed overview of the main statistical technique used throughout this project, LTA, both algebraically and conceptually. Following this, a literature review of this technique will be displayed to identify the current usage of LTA within MSK research, and relevant extensions that could be utilised.

Chapter 3: **Latent Transition Analysis**

3.1 Introduction

When faced with individuals who have varying responses to particular questions such as those regarding pain and function, it may be possible to uncover hidden (latent) sub-groups of people that have similar characteristics. Latent Class Analysis (LCA) is a "person-centred" technique (Muthén and Muthén, 2000) that defines sub-groups of people who are similar to each other and different from other sub-groups; these groups contain individuals who share similar characteristics defined by multiple factors measured at one time point. This is a frequently used method to identify clinically important sub-groups, and subsequently investigate future outcomes for these classes (such as in musculoskeletal research, pain, function or perceived health status).

Two related longitudinal approaches to this are Latent Class Growth Analysis (LCGA) and Longitudinal Latent Class Analysis (LLCA), which cluster individuals together based on one factor measured at multiple time points, to determine common patterns or trajectories of this single measure over time. These approaches can be extended by the use of Latent Transition Analysis (LTA).

LTA will be explained in more detail in *section 3.3* and *section 3.4*. However, briefly, LTA defines 'states' (instead of 'classes') based on multiple factors measured at each time point, and the probability of moving between these states over the period(s) of time analysed can be estimated. This provides additional information, based on their inclusion in a particular state at time 1, regarding the probability that they remain in the same state (or switch to another) at time 2. For more than two time points, the probability of remaining in the same

state at time 3 can be estimated, depending (conditional) on their inclusion in that state at time 2 (and similar for further follow-up time points).

3.2 Objectives

The specific objectives of this chapter were to:

- Explain the background and theory to LTA (both algebraically and in context);
- Review papers using LTA in the field of MSK conditions;
- Identify the current fields and application of LTA in current research;
- Highlight aspects of the technique that will be beneficial to the overall aims of the PhD.

3.3 Latent Transition Analysis

This section will introduce LTA in a more formal way. Algebraically, the process is made up of several steps that make up the sets of equations that are used to represent LTA (Collins and Lanza, 2010).

Participants provide data on observed questions $j = 1, ..., J$ where J represents the total number of questions assessed. These J questions (indicators) are assessed at $t = 1, ..., T$ times (i.e. number of follow-up points). The observed indicator j has a specified number of response categories, $r_{j,t} = 1, ..., R_{j,t}$ over the time points t. Here, for simplicity, the number of response options are presumed to be the same for each question and are identical across time within a question, therefore $R_{j,1} = R_{j,2} = ... = R_{j,T} = R_j$. Hence, if the question has a 'yes' (Y) or 'no' (N) answer, then $R_j = 2$. A contingency table can be formed by tabulating the J indicators at T time points which will have W cells where W is represented as

$$W = \prod_{t=1}^{T} \prod_{j=1}^{J} R_j \qquad (\mathbf{3.1})$$

So for example, if 5 questions were to be included in the LTA analysis (J= 5), at 3 time points (T= 3), and the response for each question at each time point were 'yes/ no' (i.e. R_j= 2), then $W= 2^{3x5}= 2^{15}$= 32,768 possible cells.

Within each of the W cells, there is a complete response pattern which is a vector of the responses to the J indicators at the T time points. The actual patterns of response to questions are represented by $\boldsymbol{y}= (r_{1,1}, ..., r_{J,T})$. If we let $\boldsymbol{Y}$ represent the array of response patterns, then $\boldsymbol{Y}$ contains W rows, and $T \times J$ columns. So, for a simple example where there are 2 questions (J= 2), each with response options Y or N (R_j= 2), measured at 2 time points (T= 2), the following would be the 16 potential response patterns in the array $\boldsymbol{Y}$:

$$\begin{pmatrix} NNNN \\ YNNN \\ NYNN \\ NNYN \\ NNNY \\ YYNN \\ YNYN \\ YNNY \\ NYYN \\ NYNY \\ NNYY \\ YYYN \\ YYNY \\ YNYY \\ NYYY \\ YYYY \end{pmatrix}$$

Each response pattern $\boldsymbol{y}$ has a probability $P(\boldsymbol{Y}= \boldsymbol{y})$, where $\Sigma P(\boldsymbol{Y}= \boldsymbol{y})$= 1, i.e. the probabilities of all of the combinations sums to one.

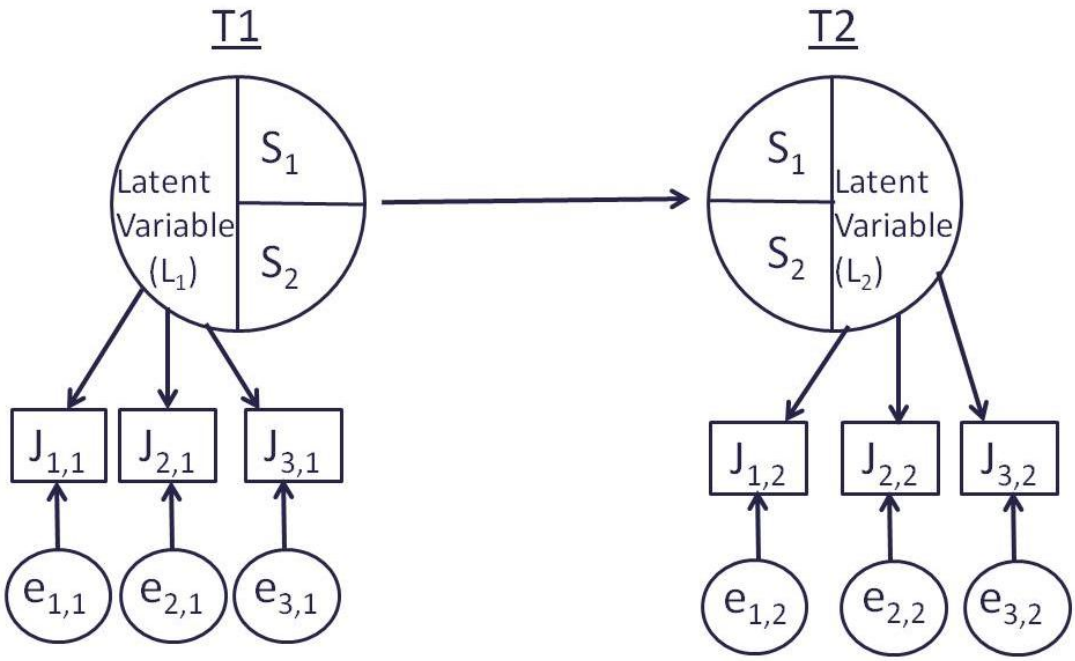
If the overall categorical latent variable is notated as L, where L has S latent states, then at time 1, L_1 is defined as s_1= 1, ..., S (any value from 1 up to S), at time 2 L_2 is defined as s_2= 1, ..., S, and so forth until L_T for time T, when s_T= 1, ..., S. Again, it is presumed that the

number (and profile) of latent states at each follow-up point is the same (which is typically the case), therefore $S_1 = S_2 = ... = S_T = S$.

All of the data at all follow-up points in a LTA model are estimated at the same time, so some states may have a low proportion of people estimated to be in that state (even potentially zero members) and therefore may not be represented at one time point. However, those states may be present at other time points in the follow-up and hence must be represented so that individuals can transition into that group at a later follow-up point. Therefore, it is possible for a state to be empty (zero prevalence) at one or more time points, but not at all time points.

Figure 3.3.1 presents a simple LTA model with 2 latent states identified (S_1 and S_2) based on 3 indicators (J_1, J_2 and J_3), measured at 2 time points (T_1 and T_2). It can be seen that there is some error (e) associated with each indicator, therefore reflecting that individuals may potentially be designated into states incorrectly. I have developed *Figure 3.3.1* based on the diagrammatic presentation of LCA in Collins and Lanza, 2010.

Figure 3.3.1: Diagrammatic representation of Latent Transition Analysis of two latent states identified from three indicators measured at two time points.



The first sets of parameters that are calculated by the LTA process are the latent state proportions, which represent the number of individuals estimated to be in each latent state at each time point. The proportion in latent state *s* at time *t* is denoted as δ_{st} (delta). The sum of all the proportions across all latent states equates to 1 for each time *t* and can be represented by equation (**3.2**) as membership of latent states are mutually exclusive and mutually exhaustive, i.e. individuals must belong to one and only one state at each time point.

$$\sum_{s_t=1}^{S} \delta_{s_t} = 1 \qquad (\mathbf{3.2})$$

The second sets of estimates are the item-response probabilities which reflect the characteristics (in terms of question responses) of the individuals in each state. These parameters $\rho_{j,r_{j,t}|s_t}$ (rho) represent the probability of response $r_{j,t}$ to the observed indicator *j*, conditional on latent state s_t at time *t*. This means, at a time point, this is the estimated probability of response to a question given an individual is a member of that particular state. The sum of these probabilities across responses within each indicator within a state is equal to 1, and represented by equation (**3.3**):

$$\sum_{r_{j,t}=1}^{R_j} \rho_{j,r_{j,t}|\, s_t} = 1 \qquad (\mathbf{3.3})$$

These probabilities are used as the basis of assigning descriptive labels to the latent states. Often these parameters are constrained to be equal across time, i.e. same item-response probabilities for each latent state at time 1 and time 2, therefore state interpretation does not change over follow-up.

The third sets of estimates are the transition probabilities (τ (tau)). We define $\tau_{s_{t+1}|\, s_t}$ as the probability of transition to latent state *s* at time *t+1*, conditional on inclusion in latent state *s* at time *t*. The probabilities are displayed in a matrix as shown below (**3.4**):

$$\begin{pmatrix} \tau_{1_{t+1}|1_t} & \boldsymbol{\tau_{2_{t+1}|1_t}} & \cdots & \tau_{S_{t+1}|1_t} \\ \tau_{1_{t+1}|2_t} & \tau_{2_{t+1}|2_t} & \cdots & \tau_{S_{t+1}|2_t} \\ \cdots & \cdots & \cdots & \cdots \\ \tau_{1_{t+1}|S_t} & \tau_{2_{t+1}|S_t} & \cdots & \tau_{S_{t+1}|S_t} \end{pmatrix} \quad \textbf{(3.4)}$$

where for example, $\tau_{2_{t+1}|1_t}$ (in bold) represents the transitional probability of membership in state 2 at time *t*+1 conditional on membership in state 1 at time *t*. Individuals can only belong to one latent state at each time point, so these probabilities sum to one for an individual at each time point:

$$\sum_{s_{t+1}=1}^{S} \tau_{s_{t+1}|s_t} = 1 \quad \textbf{(3.5)}$$

When the LTA process classifies individuals into a latent state, it generates 'posterior' probabilities that represent each participant's probability of being in each state at each time point, based on the participant's responses to the indicators. For each time point, an individual will have a probability of being placed into each state, and they are allocated to the state in which they have the highest probability of belonging.

In most of the published examples, LTA has been carried out between two time points; however, it is possible to model over more than two time points. In order to create the latent states efficiently and appropriately, all the time points must be analysed in the same model (i.e. not *t1* to *t2* and then separately *t2* to *t3*). It is usual to present the transitional probabilities from *t1* to *t2*, and then *t2* to *t3*, but the software will model the transitions in one modelling process.

3.4 Applied example of Latent Transition Analysis

The focus of this section is to describe a practical example using the principles and common estimates developed with a LTA process, using the example of a paper published by Cleveland et al., 2012.

As explained in more detail in *3.3*, there are three key parts estimated with LTA. These parts are represented by the Greek letters δ (delta), ρ (rho) and τ (tau). To help explain these estimates in terms of application, reference to the study reported by Cleveland et al., 2012 will be made. One of the main aims of this study was to investigate the drinking habits of adolescents in the United States both pre-college and then in the subsequent fall (autumn) semester of college students (approximately 18-19 years of age). The study performed a LTA using data from the baseline and follow-up assessments and produced the data shown below in *Table 3.4.1* (taken from *Table 3* in the original manuscript with permission from the journal). From the seven dichotomous items measured at the two time points, the authors determined that the optimum model contained four latent states, with the states labelled: non-drinker, weekend non-binger, weekend binger and heavy drinker.

The δ's represent the proportion of students estimated to be in a particular latent state. For example, at Time 1 (baseline), the proportion of the study population that were represented by the first latent state ('non-drinker') was 0.42, therefore 42% of the sample. As the states are mutually exclusive and mutually exhaustive (i.e. all participants must belong to one, and only one, of the defined latent states) the sum of the proportion estimates equate to 1.0 (subject to rounding error); in this particular example at Time 1 these figures are 0.42+0.20+0.30+0.08= 1.00. It can be seen from the table that at Time 2 these proportions appear to shift towards the heavier drinker; implied by the increase of the estimated proportion in the 'heavy drinker' state from 0.08 at time 1 to 0.28 at time 2. These delta parameters are a useful indication of the relative frequency of the states at various stages, however, it does not specify an individual's movement through the transition from one state to another (which is provided by the τ parameters at the bottom of *Table 3.4.1*).

Table 3.4.1: LTA estimates from Cleveland et al., 2012 representing the alcohol drinking habits of adolescents (reproduced with permission from the American Psychology Association).

	Latent state			
	Non-drinker	Weekend non-binger	Weekend binger	Heavy drinker
Latent state proportion (δ):				
Time 1 (Baseline)	0.42	0.20	0.30	0.08
Time 2 (Fall follow-up)	0.30	0.19	0.22	0.28
Item-response probabilities (ρ)				
Drink in past month	0.10	0.99	1.00	1.00
Been drunk in past month	0.00	0.53	0.99	1.00
Weekday (Sun, Mon, Tues, Weds) drinking	0.00	0.05	0.08	0.30
Thursday drinking	0.00	0.06	0.02	0.74
Weekend (Fri, Sat) drinking	0.02	0.65	0.88	0.98
Binge drank in past 2 weeks	0.00	0.10	0.83	0.95
Peak BAC>0.08	0.00	0.23	0.89	0.94
Latent transition probabilities (τ) T1(rows) to T2 (columns)				
Non-drinker	**0.64**	0.21	0.11	0.04
Weekend non-binger	*0.14*	**0.43**	0.24	0.18
Weekend binger	0.02	0.06	**0.41**	0.51
Heavy drinker	0.01	0.03	0.00	**0.96**

Footnote: BAC= Blood Alcohol Content; bold entries represent stability.

The ρ's represent the item-response probabilities. These probabilities reflect the estimated indicator response for individuals classified within each latent state. Focussing on the 'weekend drinking' question, it can be seen that the item-response probabilities across the latent states suggest that participants in the 'non-drinker' state had a very low probability (0.02) of responding 'yes' to drinking at the weekend. However, the analogous item-response probability for the other states ('weekend non-binger', 'weekend binger' and 'heavy drinker') contained high probabilities of answering 'yes' to this question (0.65, 0.88 and 0.98 respectively).

The final section of the table presents the transition probabilities (τ). These are the estimated probabilities that an individual will subsequently transition to (or remain in) a particular latent state, conditional on previous state membership. For example, the value of

0.14 (in italics in the table) indicates that there is a probability of 0.14 that a participant in the 'weekend non-binger' state at Time 1 will subsequently move into the 'non-drinker' state at Time 2. The four probabilities on the leading diagonal in bold represent stability, i.e. the probability that an individual is estimated to be in the same latent state at Time 2 conditional on the same state membership at Time 1. The highest of these values is 0.96, meaning that there is a probability of 0.96 for remaining in the 'heavy drinker' state at Time 2 given membership in the 'heavy drinker' state at Time 1. The combined probabilities for transitioning out of the 'heavy drinker' state from Time 1 to Time 2 is very small at 0.04: 0.03 chance of transitioning to 'weekend non-binger', 0.01 to 'non-drinker' and <0.001 to 'weekend binger'.

The next section reviews the fields in which LTA has been used. This will aid understanding of the extensions and limitations to the technique, and explore the extent to which LTA has previously been used in MSK research.

3.5 Literature search strategy

A systematic database search was performed in MEDLINE (Medical Literature Analysis and Retrieval System Online) (through NHS National Library for Health) of manuscripts published in English between 1966 (the earliest date for papers within MEDLINE) and September 2015.

3.5.1 Search term

The search term used was:

"Latent Transition Analysis" in the title or abstract of the paper.

A decision was made to not include the abbreviation of the technique, i.e. LTA, as this abbreviation/ acronym has several meanings that could be relevant in a health manuscript

search, such as ‘Life Threatening Allergy’, ‘Leisure Time Activity’ or ‘Late Term Abortion’, in addition to the technique of interest.

Searching of the term “Latent Transition Analysis” was restricted to the title of the manuscript and/ or inclusion in the abstract as it was felt that if LTA was a key technique applied in the study, it would be highly likely that the term would be referenced in at least one of these two areas of the paper. Not using the term “Latent Transition Analysis” as a search term for title or abstract would result in a very large number of hits, where full papers would need to be retrieved to determine the analysis method used.

3.5.2 Inclusion/ exclusion criteria

Inclusion criteria were:

i. Those using/ describing the technique of LTA;
ii. Those with at least the abstract in the English language.

Exclusion criteria were:

i. Those that made reference to the method of LTA, but did not show any evidence of actually using the technique.

For this review the manuscripts to be included needed to be in the English language, either in their published form or after translation. As it was not feasible to translate all non-English language manuscripts returned by the application of the search term, a decision was made to restrict inclusion to those manuscripts with an English language abstract and by assessing the abstract, it could be decided if the full paper needed to be retrieved and translated.

Manuscripts that made some reference to the Latent Transitional Analysis technique but did not use the method in the manuscript were excluded. Examples of where this occurred

were references to the application of LTA presented in another manuscript or reference to the technique as an alternative methodology.

In addition to the automated database search, references from all relevant papers found in the search were checked to ensure any potentially useful and relevant manuscripts were not missed by using the specific research criteria.

As a result of these reference checks (presented below) it became apparent that some relevant papers had not been picked up from the original search and hence an additional search in the 'Web of Knowledge' was carried out to ensure all potential papers were identified. (N.B. The 'Web of Knowledge' was used here to allow searching of an additional database, and for the citation analysis, explained below). This included two additional steps; firstly, the search terms were expanded to include "Latent Transition Model" or "Latent Transition Analysis" anywhere in the title and/ or abstract. This resulted in the inclusion/ exclusion criteria being modified to include the requirement of 'those using/ describing the technique of Latent Transition Analysis OR Latent Transition Model'. The other inclusion/ exclusion criteria remained the same.

The final step involved a citation analysis on the two main original manuscripts describing the development of the LTA methodology, namely Graham et al., 1991 and Collins and Wugalter, 1992. The modified inclusion/ exclusion criteria were also applied to papers gained through the citation analysis.

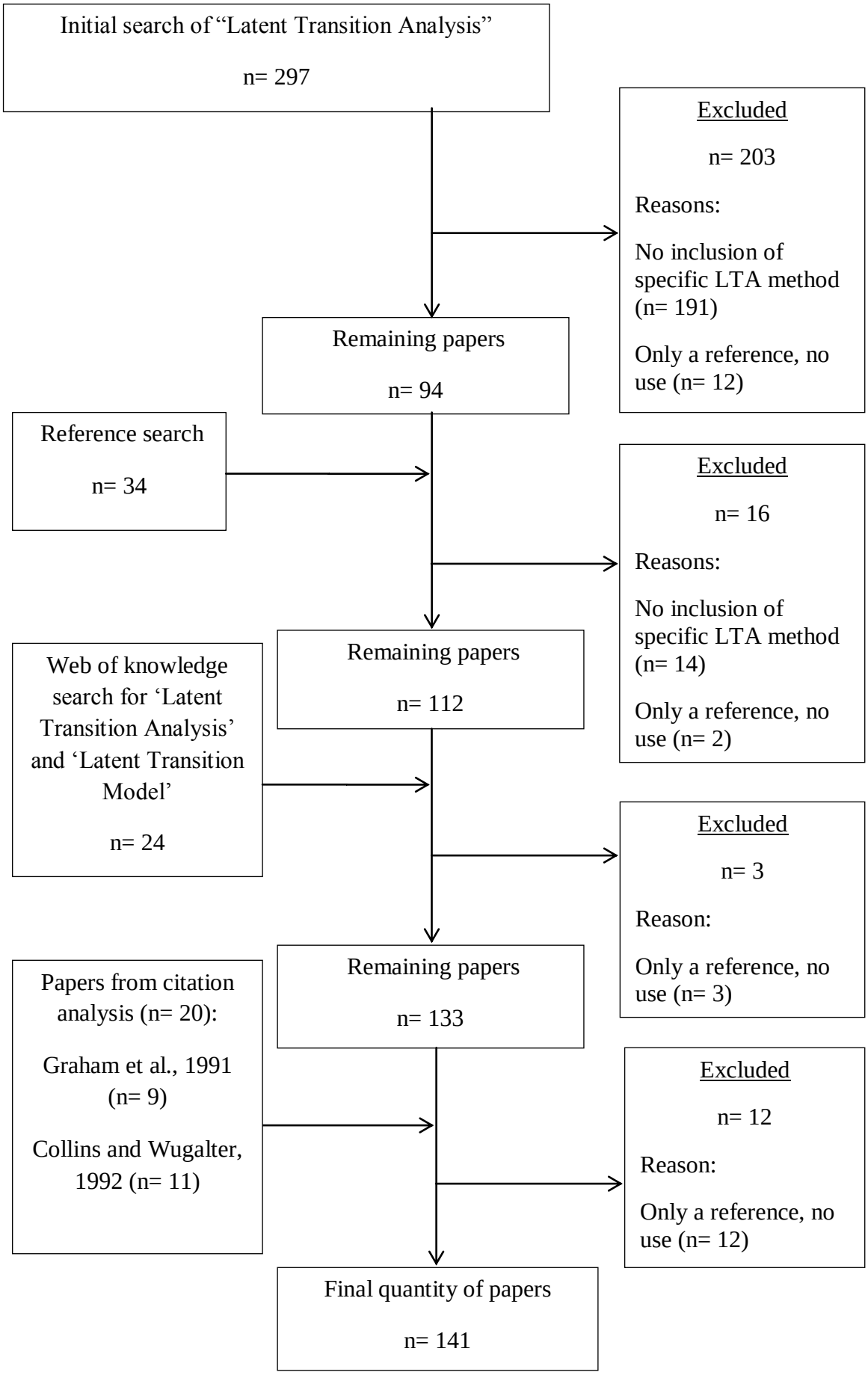
3.6 Application of search strategy and inclusion/ exclusion criteria

The process of the inclusion of papers is presented in *Figure 3.6.1*. This flow diagram indicates the specific sequence of how the papers were included in the literature search, as

well as those that were excluded with the relevant reasons. In summary, 297 papers were initially identified from the search in MEDLINE, of which 203 were excluded as they did not use the method of LTA; the majority of these were picked up as the words ‘latent’, ‘transition’ and ‘analysis’ appeared in the title or abstract, but the words were actually used to refer to other areas, frequently genetics.

After checking references, the majority of these exclusion decisions were made due to the fact that the manuscripts only make reference to the technique and did not actively use LTA; as the main focus of the literature review was to identify ways in which the technique has been used and extended, which resulted in their exclusion. After the searches and citation analyses were completed, 141 papers met the criteria and made up the final number of relevant papers.

Figure 3.6.1: Flow diagram of manuscripts included in literature search.



3.7 Fields applied to Latent Transition Analysis

3.7.1 Health

Of the 141 papers identified, 96 (68%) were categorised as health related. The main reoccurring topics within health were substance use (24 papers), alcohol consumption (20 papers), and psychological disorders (such as eating disorders, depression and anxiety) (16 papers). These papers generally expressed the standard use of LTA by assessing individuals at multiple time points and investigating any particular change in latent state at two or more time points. Less frequent application was in smoking cessation (n= 12), obesity/ physical activity (n= 5) and sexual health (n= 4). Only two applications of LTA to MSK conditions were found, (Reboussin et al., 1999; Von Korff and Miglioretti, 2005) which are described in *section 3.8*.

3.7.2 Education

35 papers were from the field of educational research; these consisted of 15 papers investigating topics related to adolescence and child personality development (papers including adolescence and alcohol, for example, were included under the health category), eight papers focussed on child educational development (such as reading and arithmetic), two on bullying and two on racial discrimination patterns. Moving from health to education did not alter the application of LTA, with the majority of manuscripts presenting a standard application to determine parameter estimates. Some researchers presented extensions to the standard process that is discussed in subsequent sections.

3.7.3 Statistical modelling

A further 10 papers in the review were related to advances in the statistical modelling of the LTA method. These papers spanned a broad period of time from the seminal work of Collins and Wugalter in 1992 applying the technique to the mathematical skill of students,

to Chung et al. in 2008 who investigated the impact of performing a simulation study with LTA in small samples (Collins and Wugalter, 1992; Chung et al., 2008). The papers in this section also included work that offered key insights into the flexibility of LTA when faced with certain issues such as missing data and how LTA can be extended to model two latent variables simultaneously, i.e. ALTA; these are described in detail later in this chapter (*sections 3.9.5* and *3.9.6*, respectively).

3.8 Latent Transition Analysis in musculoskeletal research

The purpose of this section is to present the current, published applications of LTA within the field of MSK research with the purpose of providing a snapshot of how much this technique has been utilised.

3.8.1 Published examples

The search identified only two papers that had used LTA in the field of musculoskeletal disorders; Reboussin et al., 1999, and Von Korff and Miglioretti, 2005.

In the first paper by Reboussin et al., 1999, estimating equations were combined with the LTA approach, and therefore the probability of transitioning into a latent state at a follow-up time period is modelled by a logistic regression for nominal responses (Reboussin et al., 1999). The paper used data from the Longitudinal Study Of Aging, specifically six self-reported items (walking a quarter mile, climbing 10 steps, standing two hours, stooping/ crouching/ kneeling, reaching up over head and lifting 25 pounds) which all had a binary (yes/ no) response. The aim of the paper was to investigate physical disability in older people over 6 years, predicted by age, gender, and presence/ absence of arthritis at baseline.

The authors identified three states labelled ‘no disability’, ‘moderate disability’ and ‘marked disability’ (high probability of disability) and found that the majority of the transitions were from ‘no disability’ to ‘moderate disability’ (identified using logistic regression rather than through a transition probability matrix). The authors then included factors of interest into the modelling procedure (i.e. age, gender, presence of arthritis) in order to predict state membership at the follow-up time point. The applied example provided by the authors demonstrated that an 86 year old female with arthritis at baseline had a 37.4% chance of transitioning from the ‘no disability’ state to the ‘moderate disability’ state over the 6 years. However, a male of the same age without arthritis at baseline had a 12.9% chance of the same transitional pattern.

While this was displayed as a different modelling procedure (using logistic regression to predict future state membership, and not estimating transition probabilities), there are distinct parallels with that of the LTA process, and using covariates to predict movement (discussed later in *section 3.9.3*). In terms of the application within MSK, the detail was unfortunately limited, mainly due to the fact that this was not the main focus of the paper; due to the high mathematical content, it was clear the main focus of the paper was to present an alternative methodology.

The second paper that investigated MSK with a LTA-related model was Von Korff and Miglioretti, 2005. It is worth noting that the modelling technique used in this paper is published in more detail in Miglioretti, 2003; however, the focus of the 2005 paper includes more detail on the application in people with back pain, so this is the manuscript that is discussed.

The authors apply Latent Transition Regression Analysis (LTRA) which is a Bayesian approach to LTA, and permits the opportunity to jointly analyse a mixture of longitudinal

outcomes (not just categorical), from any distribution (this is the Bayesian aspect). Bayesian analysis in statistics incorporates the position that prior knowledge is available on the research topic and that the findings from previous research on similar questions can help determine a particular distribution of the data. By analysing data with a Bayesian framework, it is possible to specify the distribution of the data along with setting the starting values for estimates, i.e. from the prior information.

The main objective of this paper was to develop a prognostic method for assessing back pain using data on 18 to 75 year olds from primary care visits in a health care system (Group Health Cooperative) in Washington State. The participants were assessed at baseline, 1 and 2 years and, through the use of LTRA, were divided into four classes (i.e. latent states) - 'no pain', 'mild pain', 'moderate pain and limitation' and 'severe pain and limitation'. These classes (states) were based on two continuous measures: pain intensity and pain interference; four dichotomous items: pain impact score, unable to work, unable to work in prior year (in newly retired persons) and unable to work in prior year (excluding newly retired persons); and one ordinal item: number of disability days in prior 6 months (0, 1-6, 7-14, 15-30, 31+).

With regard to the transition probabilities, the authors found that the higher probabilities related to those who remained stable (stayed in the same state) over the two sets of time periods investigated (baseline to year 1, year 1 to year 2). The authors then investigated prognostic variables amongst patients with mild or moderate back pain. Those with mild or moderate back pain, high levels of pain persistence, diffuse pain or depressive symptoms, were not predictive of the risk of future severe limiting pain. However, in the whole analysis population, those with high depression symptoms were at higher risk of future severe back pain. Finally, the authors created a predictive risk score based on baseline pain scores, and used this to predict future back pain severity, with those expressing more

likelihood of chronic pain being more at risk of chronic back pain at 5 years (5 year data not used in the transition analysis).

3.8.2 Related methodologies

One advantage of using LTA in MSK research with longitudinal data is the potential to provide a more informative picture of how patients are progressing (in terms of improvement or deterioration). However, as seen from the results of the literature review, LTA with MSK conditions has not received much attention. Other methods, such as Longitudinal Latent Class Analysis (LLCA) and latent class growth analysis (LCGA), whereby one factor measured at multiple time points is examined to determine common patterns over time within this factor, have been utilised more (Dunn et al., 2011; Verkleij et al., 2012; Dunn et al., 2013; Holla et al., 2014; Nicholls et al., 2014; Rzewuska et al., 2015).

LLCA and LCGA have clear benefits over standard regression techniques; they can pick up on fluctuating trajectories over time (e.g. pain responses over follow-up) and can also investigate which characteristics are predictive of membership within particular classes. However, in some contexts, this is not sufficient in order to develop a more detailed picture about individuals' response over follow-up. Trajectories of health conditions are unlikely to be sufficient based on one single factor, and should therefore be modelled on multiple important and related factors. This is one area in which LTA becomes advantageous because states are defined on multiple characteristics, as well as identifying the transitions that individuals are likely to make, and provide an idea of the number of individuals who will make that transition through the transitional probabilities.

It is important to note that the technique of LTA was derived from another method, the Latent Markov (LM) model. At the time LTA was developed (Collins and Wugalter,

1992), the LM model grouped individuals into states at each time point based on one factor measured at multiple time points, and then investigated transition probabilities between follow-up points. However, since the development of the LTA, researchers have also extended the LM to assess more than one factor measured at multiple time points (whilst still exploring transition probabilities at follow-up); therefore the LM model and LTA are very similar approaches. However, there is one main difference between the approaches; the LM models assume class membership without error, therefore it is presumed individuals are perfectly assigned to their optimum class (Kaplan, 2008). Importantly, this is not assumed in the LTA process (through the use of posterior probabilities, explained in *sections 3.3* and *5.3*). Therefore in most research fields, where perfect state designation is rarely a valid assumption, LTA is arguably more suitable.

LTA incorporates data from both baseline and future follow-up points thus developing a more accurate assessment of the study period, compared to a simpler approach which might look at the baseline data separately from the follow-up data. In addition to this, potential extensions available for LTA could reveal further detailed information in the field of MSK disorders; the next section will present some of these extensions.

3.9 Relevant extensions of Latent Transition Analysis

This section presents some of the relevant applications and extensions arising from the literature review. In the first section, additions in the approach to LTA such as within sub-groups (stratified) analysis, more than two time points analysed, and parameter constraints are explained. In addition to this, further analytical extensions such as using factors to predict transitional movement/ stability and modified LTA approaches are described.

3.9.1 Additional follow-up points

In the example given in *3.4* only two follow-up time points were presented, but LTA can model many time points (Mplus can model >4 for example, Muthén and Muthén, 1998-2015). To present this in context, Cain et al., 2010 performed a study investigating patterns of eating and weight. The data was collected over 3½ years (baseline fall (autumn) 1st year, then spring in 2nd, 3rd and 4th year) and the LTA process was applied between each of the neighbouring time points (fall 1st year to spring 2nd year, spring 2nd year to spring 3rd year, and 3rd year to 4th year); these investigations formed the main section of analyses. However, subsequent to this, the authors performed a sensitivity analysis by repeating the LTA process, but this time investigating transitions just from baseline to the final follow-up point (1st year autumn to 4th year spring). If the resulting transition estimates produced different results from those found in the yearly LTA process, it could be argued that shorter follow-up periods are necessary for research in this study. In addition, the results from the sensitivity analysis would allow examination of whether the patterns change gradually at each follow-up time resulting in a greater cumulative difference over the four years, or if the conditions remain pretty much the same over a longer time period. The authors found that the pattern over the full time period were similar to the investigations over each single year period.

For the PhD project described in this thesis, reported in more detail in *section 4.3*, there were three time points available for investigation, which are used throughout the project.

3.9.2 Latent Transition Analysis within sub-groups (stratification)

One extension to the LTA approach is to perform the analysis stratified by certain pre-specified sub-groups, for example, age group or gender. These characteristics were employed by Shin where investigations into alcohol and drug consumption were analysed

(Shin, 2012). The benefit of reproducing the analysis within sub-groups (males and females, for example) is that it makes it easier to compare the transitional patterns. In the study by Shin, three latent states were found that generally represented three increasing levels of drug/ alcohol consumption (Shin, 2012). The gender stratified results indicated that between the two times points studied, females had a <0.01 probability of moving from the lowest drug intake to the highest, while males had a 0.74 probability of the same transition.

This is a useful addition and is relevant to the proposed analysis of people with hand pain/ problems as gender and age differences have already been highlighted in previous literature. This extension to the method permits the exploration of different patterns by various characteristics, which could remain unobserved in a standard LTA.

3.9.3 Prediction of transition/ stability

It is possible to include covariates in the LTA process to investigate the relation of a particular variable with the latent states. This is achieved by including pre-specified predictors in the modelling stage, and will indicate if the predictor is associated with estimating transitional movement or stability, and whether their inclusion improves model fit/ accuracy of state identification.

To present an example of this, the paper described in *3.4* (Cleveland et al., 2012) included additional analyses which investigated the effect of two different interventions focussed on alcohol monitoring (brief motivational interviews (BMIs) and parent-based interventions (PBIs)) on the transition probabilities (τ). Over the course of the study period, the students in the study were subjected to either one, both or neither of the interventions. The effect of these interventions, with reference to latent state transition/ stability, were examined by applying them as individual covariates (BMIs – yes or no; PBIs – yes or no) and as an

interaction to explore the additional effect of the students who received both: BMIs and PBIs.

The method used to analyse the importance of these covariates in predicting transitions created a series of nested models by starting with terms for each of the main effects (BMIs and PBIs) and their interaction (BMIs × PBIs). The model was then simplified by removing terms of higher order one at a time and testing for a significantly worse model fit via the Likelihood-Ratio Test (LRT) (similar to the process of model building in standard regression models).

The LRT test indicated that the transitional probabilities into the 'heavy drinker' state were different among the four treatment groups (no intervention, PBIs, BMIs, PBIs and BMIs) as the model with the main effects and the interaction was superior to the model with no covariates. The first simplification step was to remove the interaction term which yielded a non-significant result; hence the effects of the two interventions were not influenced by each other. The removal of both PBIs and BMIs led to a significant change in the LRT, indicating that students who underwent either intervention had a significantly reduced likelihood of remaining in the 'heavy drinker' state. To summarise this result, students who received either PBIs or BMIs were less likely to move to (or remain in) the 'heavy drinker' state at time 2; however, this association was not significant for the students who received both interventions.

This extension to the standard LTA process has clear benefits within the majority of investigations, and is explored in this PhD project. Important characteristics, such as age and gender, may influence the identification of the latent states, and may need to be included in the modelling process. In addition to this, gender or age group may potentially

be predictive of an improvement or more rapid deterioration in hand pain/ problems over time.

3.9.4 Parameter constraints

A slightly more technical addition to the LTA process is the ability to employ parameter restrictions to ensure certain transition pathways are not possible. A good example to illustrate this is found in Schumann et al., 2002, who investigated the readiness to quit smoking in a general population sample (aged 18-64) in a northern city of Germany over a 6 month period. The authors used the five stages in the Transtheoretical Model of Behaviour Change: precontemplation (PC), contemplation (C), preparation (P), action (A), and maintenance (M). Five nested models were hypothesised and then fitted via LTA:

1) One-stage forward movement only (no reverse transitions allowed and only one stage forward transition allowed, e.g. PC to C but not PC to P, A or M);

2) One- and two-stage forward movement only (no reverse transitions allowed and only up to two transitional advances allowed (e.g. PC to C or PC to P);

3) Forward and backward movements of one stage;

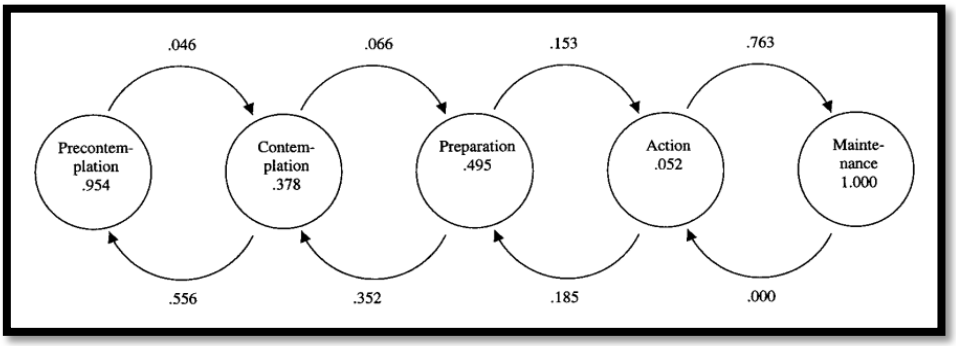
4) Forward and backward movements of up to two stages;

5) Forward and backward movements of one, two and three stages (so the only two impossible transitions in this scenario would be PC to M and vice versa).

In order to achieve these varying movement models, certain parameters were fixed to zero. Similar to the previous section (*section 3.9.3*), these models are all nested so it is possible to test for significant differences to previous models by the use of a LRT. Model 3 (forward and backward movement of one stage only) was optimum and deemed the most appropriate representation for the data. A LRT between model 3 and model 4 was tested to

examine if there was any significance in adding extra pathways. The δ and τ parameters for models 3 and 4 had very small probabilities for the additional transitional pathways unrestrained in these models, therefore implying the models were similar; thus, in the interest of choosing the most parsimonious model, model 3 was selected as the most appropriate for the 6 month follow-up period. The τ probabilities are displayed in *Figure 3.9.1* below, highlighting the model approach of a one-stage forward and backwards model.

Figure 3.9.1: Illustration of transitional probabilities (τ's) for model 3 defined in Schumann et al., 2002 (reproduced with permission from Preventive Medicine).



Another potential reason for the use of parameter restrictions might be related to the design of the study. For example, part of the analysis might presume that once you have gained a skill/ knowledge/ milestone, you can never lose it (this is also known as the Guttman process). For example, Patrick et al., 2009 investigated substance use, and one of their latent states described those who had ever consumed alcohol in their lifetime; by definition, once you have ever consumed alcohol, you can never 'untake' it, so the transitional pattern of going from alcohol use to never is illogical. To incorporate this into the model, this transition path was classified to be zero and was not estimated in the model.

A further reason why a transitional pathway might be classified as zero could be because of low frequency. In some occasions, if the numbers of people transitioning over a certain pathway is low, the software can struggle to estimate it (the model could struggle to

converge), therefore a necessary step could be to restrict that entry to zero to aid convergence. In this analysis, should convergence in LTA be an issue, potential parameter restrictions will be considered.

3.9.5 Missing data

An issue prominent in all research is the problem of missing data. Missing data is included in the simulation process of LTA (and Latent-related methodology in Mplus) by estimating parameters with maximum likelihood estimation using the EM algorithm; this is a technique used to estimate parameters based on the available data. One important issue to note is that data is presumed to be missing at random (MAR). One minor drawback is that coefficients used to predict transitions (such as gender or age) cannot contain missing data; therefore, imputation of missing values in these covariates will need to be computed separately, or observations with missing data will be excluded during the computation process.

There are two clear reasons for using the EM algorithm to account for missing data over a complete case analysis. Firstly, deleting observations can lead to bias, the magnitude of which depends on the proportion of missing data and how much the data appears to be truly MAR. Using this algorithm can preserve the underlying characteristics of the population without totally relying on the completeness of the data. Secondly, using the algorithm can lead to maintenance of statistical power as no partial loss of data is incurred.

To investigate the impact of missing data, Hyatt and Collins explored how varying particular factors included in the LTA process adjusted the results and power of the study (Hyatt and Collins, 1998). The factors they focussed on consisted of sample size, strength of measurement parameters (ρ's), type of missingness and amount of missing data. In order

to assess these factors, numerous simulation studies were carried out with particular stringent conditions.

The main results from this investigation indicated that LTA models containing missing data will function well when the ρ parameters are strong (i.e. such that the latent states identified were distinct and item-response probabilities in the region of <0.1 or >0.9) and when the sample size is large (n>1000). In addition, even when missing data (either MAR or MCAR (Missing Completely At Random)) amounts to 50%, or sample size was smaller than 1000, strong ρ parameters are often enough to provide unbiased results. This highlights the importance of identifying a model with distinct latent states. Unfortunately, due to software issues, it was not possible for the authors to assess data that was NMAR (Not Missing At Random). As missing data occurred in multiple areas of the PhD project, the EM algorithm (default approach to missing data for latent analysis in the Mplus software) was used in the analysis of this project to avoid complete case analyses.

3.9.6 Associative Latent Transition Analysis (ALTA)

In the examples presented so far, one latent structure over time has been presented (a related number of categorical variables explored to find a certain number of latent states); however one development to this, first published by Flaherty, 2008, explains how to identify two latent structures, and investigate them over time simultaneously, called Associative Latent Transition Analysis (ALTA) (Flaherty, 2008). In the published example, the two separate structures identified were based on psychological state and substance use assessed over a 1 year follow-up in US students aged 12 to 17 years old. The investigation into psychological state revealed three latent states ('positive', 'neither' and 'negative') and substance use identified six latent states ('no use', 'tried alcohol', 'tried marijuana', 'past 12 month alcohol use', 'alcohol with drunk/ binge' and 'drunk and

marijuana'). These two separate structures could be followed up individually to see how the students change in each domain, however by employing the ALTA method, it is possible to see how the two conditions relate to each other over time.

The ALTA process has three main sets of probability estimates; the first is the β (beta) probabilities which, in this example, represent the association of the substance use classes at time 1 (T1), conditional on psychological state at T1. The results indicated that at this assessment, individuals that were classified as being in a 'positive' psychological state were more likely to be in the 'no use' (0.30) or 'tried marijuana' (0.44) states and have very low probabilities of being in the higher substance abuse states. In contrast, people in the 'negative' psychological state had a low likelihood of reporting no substance use and had a high likelihood of belonging to the 'drunk and marijuana' state (0.25).

The second sets of probabilities are defined as the ε (epsilon) estimates. These denote the psychological state at T2, conditional on latent class memberships for substance use and psychological state at T1, with the data represented in three tables, one for each psychological state ('negative', 'positive', 'neither') at T1. Some important results from this indicated that reporting no substance abuse at T1 was strongly related to stability in psychological state for those 'positive' at T1, i.e. 0.76 remained in a 'positive' state at T2, whereas stability was lower for those in the 'drunk and marijuana' state at T1 with 0.41 chance of remaining in the 'positive' psychological state at both time points. When investigating those in the 'negative' state at T1 across all the six substance classes, at least 0.66 remained in the 'negative' state at T2 (ε range 0.66-0.75). Substance use at T1 was related to transitioning from a 'positive' to a 'negative' state; more than 40% transitioned into the 'negative' state at T2 if they were in the 'past 12 month alcohol use', 'alcohol with drunk/ binge' or 'drunk and marijuana' states at T1.

The third set of results from the ALTA are represented by the η (eta) which represent the probabilities of substance use at T2 conditional on previous and concurrent levels of psychological state and the previous level of substance abuse. In this example, there are 18 tables (three psychological states ×six substance use states) representing all the possible associations between T1 and T2 states; the manuscript presents a subset of the tables. The author was able to show that those who reported 'no use' at T1, and remained 'positive' psychologically at both T1 and T2, were highly likely to remain 'no use' at T2 (0.72). Of the three sub-tables illustrated in the manuscript, the highest estimates within each were related to stability in both psychological state and substance use (η range 0.63 to 0.72).

The uses of the ALTA method could have a valid use in MSK research, as many individuals are often not affected by one MSK condition alone, or can also suffer from other comorbid conditions. Therefore, investigating how these conditions interact and develop over a period of follow-up could reveal interesting relationships. ALTA is explained in more detail in *Chapter 9*.

3.9.7 Latent Transition Analysis with GMM

One of the extensions developed by researchers is the Latent Transition Growth Mixture Model (LT-GMM) which was published by Petras et al., 2011. The key difference in this approach is that while LTA typically uses observed categorical indicators to define the latent variable (and thus latent states), the latent variable in a LT-GMM is categorised by a continuous longitudinal process during a specific time period. Therefore, individuals within each state will have a similar cluster (growth trajectory) to each other, and be distinctly different to individuals in another state. Trajectories were identified at two periods of time (grades one to three, and grades six to 12 in Petras et al., 2011) based on multiple indicators measured at multiple time points within each time period. Therefore,

individuals were permitted to be a member of one trajectory at one time period, but then transition to another at the second time period (thus permitting the calculation of transition probabilities similar to a standard LTA).

The example given in the manuscript is that of aggressive and disruptive behaviour in young students in the United States, aged between six and 18; students were assessed by their respective teacher on a relevant numerical scale between grades one and three (six to nine years of age), and between grades six and 12 (12 to 18 years of age) (Petras et al., 2011). In order to assess the data without gender bias, the approach was done for males and females separately. The investigation revealed a two-class solution for both males and females; states within each gender represented a 'high' aggression and 'low' aggression stage. Therefore, it was possible to explore whether individuals remained in a similar growth pattern in their teenage years as they did in their previous assessment. The authors then explored this technique further by splitting the transition probabilities by particular interventions that had been carried out between baseline and the follow-up period (as well as gender, mentioned previously). The authors found that males with a family-centred intervention (attempts to improve behaviour from key family members) were more likely to improve their behaviour over the study period; while females were more likely to benefit from a classroom-centred intervention (attempts to improve behaviour from teachers) (Petras et al., 2011).

This is a useful addition to the technique of LTA, as this offers an alternative approach to LCGA (of which there are parallels) when the research question aims to explore whether growth patterns remain similar over two different time periods. However, this approach is still to be researched further and potential limitations (such as model identification and complexity of software) are yet to be accomplished. At the time of this PhD project,

relevant software was not available to employ this approach, and therefore this extension was not considered for use in this project.

3.9.8 LTA-MRM

A combination method of LTA and a Mixture Rasch Model (MRM) was developed by Cho et al., 2010; however the more detailed applied example is published in Cho et al., 2011. The main distinctive difference between LTA and the process of LTA-MRM is that in the latter the data used to identify the latent states are developed using a Rasch model, rather than the observed binary data. The data in the example stems from students being tested on 14 items at a number of time points, of which each of the items relate to a particular task based on specific cognitive skills. Each of the items have a binary outcome (the student can either do the task or not). The results of the 14 items were then analysed to identify the number of underlying latent states. In this example, two latent states were sufficient to represent the data; one which indicated the students coped well with all the tasks and another which indicated that students struggled with six of the items. Then the latent states were followed up over time in a similar process to the standard LTA process (investigating the proportion of transitions from the group of students that struggled with some items, into the state of those who coped well with all tasks).

3.9.9 LPTA

It is worth commenting on an additional paper published in 2011 by Thompson et al. who devised a similar technique to LTA that was called Latent Profile Transition Analysis (LPTA) (Thompson et al., 2011). The authors state in their article that Latent Profile Analysis (LPA) uses continuous data as indicators, whereas LCA generally uses categorical data (Thompson et al., 2011). Therefore LPA identifies ‘profiles’ of responses

compared to ‘classes’. In the article by Thompson and colleagues, LPA is extended to incorporate the transition probabilities, and therefore develop LPTA.

While the extensions described above in *sections 3.9.7* to *3.9.9* are important extensions to the LTA method, the availability of software to complete these approaches is not widely available, and challenges such as optimal model identification remain. In addition to this, the added complexities of these methods are not required to complete the PhD objectives (*section 1.4*). Therefore the previous methods and extensions highlighted in this chapter were more suited to the research objectives and ‘combination’ LTA methods were not explored further.

3.10 Summary

Over *sections 3.3* to *3.9*, the standard uses of LTA, the current areas of literature in which LTA has been published and some of the most useful extensions to the technique have been reported. Notably these include performing the LTA process stratified by key variables, using predictors to investigate which factors are associated with transitioning to (or remaining in) particular states, and the extension of ALTA, permitting the investigation of the longitudinal relations between two (hypothetically-linked) latent variables in the same model. These extensions are utilised at various stages in this project.

This chapter has described the technique of LTA which is used in this PhD project. In addition to this, useful extensions to the method have been presented, highlighting those most relevant to MSK research and those to be used in this project. The next chapter details the key information collected in the NorStOP study, and develops a base model using LTA, which is used throughout the PhD project.

Chapter 4: **Development of the base model**

4.1 Introduction

This chapter describes the application of LTA to determine common profiles of hand pain/ problems in people aged 50 years and older, using key characteristics of hand pain and function identified from patient opinion and previous literature. The development of the base model and latent states of hand pain/ problems is described. The overall focus of the analysis in this chapter is to identify states of hand osteoarthritis pain/ problems, then in *Chapter 5*, investigate how these states develop and/ or change over time.

The chapter also describes the process of gaining the input and experiences from people with hand pain/ problems/ osteoarthritis, and narrowing down a large number of individual indicators related to hand pain/ problems into a model with only a limited set of indicators. These indicators were then used to derive latent states related to hand pain/ problems of the NorStOP study population of over 5,000 people aged 50 years and older, over a 6 year period. The term 'base model' is used throughout this chapter (and subsequent chapters); the model is labelled 'base' because it is the initial model to which further extensions are applied from *Chapter 5*.

4.2 Objectives

The specific objectives of the study included in this chapter were to:

i. Identify hand indicators that were important to people with hand pain/ problems/ OA (potential indicators for the analysis);

ii. Determine distinct phenotypes of people with hand OA using LTA based on the selected indicators.

4.3 NorStOP study

This section describes the study that was used for this PhD project and explains the recruitment stages for that study. An overview of the content of the questionnaires used is also provided.

4.3.1 Study design

The North Staffordshire Osteoarthritis Project (NorStOP) is a set of three population-based cohorts who completed a two stage postal questionnaire at each of the time points (baseline, 3 years and 6 years follow-up). The primary objectives of the NorStOP study were to obtain information on the prevalence and associated characteristics of joint pain in the older adult general population and investigate how the existence and severity of joint pain changed over the study period. The initial baseline surveys took place between 2002 and 2005.

When originally contacted, all participants were aged 50 years and over, and registered at one of eight general practices within North Staffordshire. Approximately 98% of the entire population of the UK are registered at a general practice, therefore, acquiring subjects through this method has clear validity when arguing a representative sample (Bowling, 1997). The remaining 2% of the population are registered at private health centres or are not registered anywhere. The study protocol is described in detail in Thomas et al., 2004.

At baseline, all registered patients at these eight practices who were aged 50 years and over were sent a postal Health Survey (HS) that addressed various demographic factors, which is explained in more detail in the next section (*4.3.1.1*). The HS also included questions

related to hand, hip, foot and knee problems; more specifically, asking if the participant experienced pain and/ or problems in any of the four locations over the previous 12 months. In addition to this, patients were asked for consent for further contact and access to medical records. If a participant agreed to further contact, and answered at least one of the location questions affirmatively, they were then sent a further postal questionnaire, the Regional Pain Survey (RPS).

The content of the RPS is explained in more detail in *section 4.3.1.2*; however, in brief, the questionnaire was broken into four sections that represented the four locations defined in the HS (hand, hip, foot and knee) and participants were encouraged to fill out all the sections for regions for which they reported pain or problems.

This process formed the 'baseline' level of the study and the 3 year process was similar. At 3 years, patients that had consented to further contact at baseline were sent a HS similar to the one sent at baseline. This questionnaire included the same joint location questionnaire inquiring about pain/ problems in the individual locations over the past 12 months, where an affirmative answer led to them being sent a RPS. The process was repeated again at 6 years. In the case of non-response at any of the stages, a reminder postcard was sent after 2 and 4 weeks to encourage return of the questionnaire. For the analysis reported here, the participants from the three NorStOP cohorts were combined and analysed together to provide a larger sample. Ethical approval was acquired for NorStOP 1, 2 and 3 separately from the North Staffordshire Local Research Ethics Committee (NS-LREC 1351 (for NorStOP 1) and 1430 (For NorStOP 2 and 3)). I (DG) was not involved in the data collection of this cohort study, however I have used the data for the analysis reported in this thesis.

4.3.1.1 Health Survey (HS)

The HS requested information regarding a participant's lifestyle, presence of pain, and physical, emotional, and social functioning, using existing and, where available, validated assessment scales. The questionnaire was divided into 16 sections, which included general demographic information such as age, gender, marital status, living arrangements, employment status, weight and height (and therefore BMI), education, qualifications and ethnicity. Other sections included: perceived health status; access to services (for example, GP); treatments used for pain; and symptoms of anxiety and depression (Hospital Anxiety and Depression Scale (HADS) (Zigmond and Snaith, 1983)).

4.3.1.2 Regional Pain Survey (RPS)

The RPS was split into four main sections which represented the four joint locations of interest (hand, hip, foot and knee) and participants were encouraged to only complete the sections that were relevant to them (based on the location specific pain questions in the HS). Each location section had a generally comparable set of questions, based on duration, location and impact of pain followed by further detailed questions relevant to each location.

The hand section included questions on:

i. Characteristics of the hand problem (duration, which hand, impact of hand problem);
ii. Hand pain, symptoms and physical features using the AIMS2 (described in *3.3.2*), AUSCAN (described in *3.3.3*), and a diagram to indicate nodes and swellings;
iii. Function difficulties (AIMS2 and AUSCAN);
iv. Health care related to hand problem (AIMS2, treatments, GP consultation, operation);

v. Self-rated impact of pastimes and hobbies;

vi. Impact of symptoms (AIMS2).

4.3.2 The AIMS2

The AIMS2 (Arthritis Impact Measurement Scales) questionnaire was included in the RPS and is one of the most widely used measures to assess the impact of hand and wrist conditions (Meenan et al., 1992). The original questionnaire spans 14 pages, and contains 78 items, however, only 16 of these items were included in the RPS. The items were split into three sections of 'Hand and Finger function' (five items), 'Arm function' (five items), 'Arthritis Pain' (five items) and 'Overall Arthritis Impact' (one items). 11 of the items were originally considered for this project (all but the 'Arm function' items), of which seven were eventually considered for use within the project. The removal of four items was due to overlap between the AIMS2 and AUSCAN questionnaires, where, because of the methods used to develop the AUSCAN (described in *4.3.3*), the AUSCAN items were preferred. Internal consistency of the AIMS-2 questionnaire has been shown to be good, with coefficients ranging from 0.74 to 0.96 in patients assessed with hand OA (Meenan et al., 1992). In addition, test-retest estimates ranged between 0.78 and 0.94, indicating good reliability (Meenan et al., 1992). In general, the questionnaire provided good psychometric properties, sufficiently able to assess hand OA.

4.3.3 The AUSCAN

There are some parallels between the AIMS2 and AUSCAN (Australian/ Canadian Hand Osteoarthritis Index). The AUSCAN was developed after, and influenced by, the AIMS2 (hence, one of the reasons for using AUSCAN items over AIMS2 for similar items). The AUSCAN tool was developed using patient interviews and clinical professionals, and items more commonly mentioned by patients and clinicians were likely to be included in

the final questionnaire. The AUSCAN represents aspects that are included in AIMS2 (in the hand and finger function section) as well as incorporating further items of pain and stiffness.

The AUSCAN is a 15-item self-reported questionnaire scaled on a five-point ordinal scale (None, Mild, Moderate, Severe and Extreme) (Bellamy et al., 2002a). The assessment tool included items related to severity of pain, limitations in function, and stiffness, and was originally developed for patients with hand OA. The 15 items cover three areas, namely pain (five items), function (nine items) and stiffness (one item).

The questionnaire has been found to have reasonable clinimetric properties (Bellamy et al., 2002b). Including patients with hand OA in the development of the questionnaire significantly improves its face and content validity. In order to assess construct validity, authors outside of the AUSCAN development process compared its subscales with more established (however, more time consuming) measures of hand conditions (Allen et al., 2006b). In a group of nearly 900 participants with hand OA, the authors firstly explored internal consistency and found that it was high for the overall AUSCAN scale (Cronbach's α= 0.96) and for the pain and function subscales (α= 0.96 and 0.94 respectively). Construct validity was demonstrated by investigating correlations of grip (Jamar hydraulic hand dynamometer), pinch strength (Jamar hydraulic pinch gauge), and a single-item pain measure with the AUSCAN subscales. It was found that the function subscale had the strongest correlation (Spearman) with grip and pinch strength (r= -0.28 to -0.12, all p<0.02), and the pain subscale had the strongest correlation (Pearson) with the single-item pain measure (r= 0.30 to 0.33 for both left and right hand, all p<0.001). It is worth noting here, that although the correlations were significant in the desired areas, the associations were weak to modest.

Further investigations by these authors found additional clinical relevance of the measures; for every unit increase in the AUSCAN function score (so poorer function), the hand grip strength decreased by 1.31kg (which is equivalent to approximately 2.5% of mean peak grip strength). For the equivalent in the pinch strength test, a unit increase in AUSCAN function score led to a decrease of 0.19kg (equivalent to approximately 1.5% mean peak pinch strength). In summary, these findings show that even a relatively small change in the AUSCAN function subscale corresponds to strength changes in older adults that may impact on a variety of daily activities.

The test properties reported in the previous paragraphs promote the use of the AUSCAN as a suitable tool for measuring pain and function in older populations with hand pain/ problems/ OA. These properties provide reassurance that constructs of hand osteoarthritis are represented by the items included in the AUSCAN, and all of the individual items were considered as potential indicators. A full copy of the AUSCAN can be found in *Appendix A*.

The remaining indicators used in this project but not described in *4.3.2* and *4.3.3* (18 indicators) were standalone questions that were developed by the NorStOP research team and have been clearly identified as such in *Appendix B* which contains the full list of indicators considered.

4.4 Methods: Phase I: Identification of indicators important to people with hand OA

The objective of this phase of the project was to determine which indicators in the data collected in the NorStOP study were the most relevant/ applicable to people with hand pain/ problems, and hence to consider as potential indicators when identifying states of

hand pain/ problems. This used a parallel process of a review of current literature and obtaining the perspectives of people with hand OA, in order to arrive at a set of indicators to be used in the modelling process.

As is clear from *4.3*, information regarding hand pain/ problems available for participants in NorStOP was very rich but with the potential for some indicators to overlap and some to have a low prevalence (for example, frequency of previous operation on hand could plausibly be low). Also, it is rarely desirable to have a very high number of indicators in an LTA model (Vermunt and Magidson, 2003; Collins and Lanza, 2010; Núñez et al., 2011). A larger number, whilst potentially revealing more information regarding different characteristics of the population, will often cause the states to be ill-defined with a higher frequency of item-response probabilities in the 0.4 to 0.6 range (discussed in section 4.6.4.4 and section 10.5). Increasing the number of indicators, may also lead to redundancy as some indicators may assess a similar aspect and therefore inclusion in the model would not reflect a benefit. It is statistically inefficient to use indicators in the model that do not improve model distinction, and therefore it is better to determine a more parsimonious model. This is similar to the process of including terms in a multivariable model, where variables that do not provide significant additional benefit to the model would not be included. Therefore, it was important to reduce the number of potential indicators to be considered in the base model by excluding the indicators that were not necessary, but still sufficiently representing the important characteristics of hand pain and function.

4.4.1 Initial set of indicators

It was not conceivable to collect additional data within the scope of this PhD project, so therefore identification of potential indicators to be used in the model was restricted to the information already collected in NorStOP. 40 individual indicators within the NorStOP

questionnaires were initially considered as potentially relevant for the identification of hand pain/ problem phenotypes. Indicators from questionnaires measuring psychological aspects, such as symptoms of anxiety and depression, were not considered as potential indicators because the phenotypes to be produced in this project were aimed to focus on aspects of pain and function. Additional information (demographic, psychological, and treatment information) are used later in this thesis to see how changes in these factors influence phenotype membership at follow-up. The 40 indicators initially identified were from the AUSCAN (15 items), AIMS2 (seven items), stand-alone questions developed by NorStOP research team (12 items), previous hand experiences (five items) and presence of nodes (one item). A full list of these indicators can be seen in *Appendix B*.

4.4.2 Key literature

The key indicators considered to reflect aspects associated with hand pain/ problems were identified from previous literature (such as Kwok's 2013 PhD thesis "Clinical aspects of hand osteoarthritis: are erosions of importance?" and Dahaghin's 2005 PhD thesis "Hand Osteoarthritis: Epidemiology and clinical consequences", as well as reviews including factors related to hand OA onset and progression, Kalichman and Hernández-Molina, 2010; Nicholls et al., 2012; Leung et al., 2013). This acted as a check to ensure indicators not highlighted by the RUG (*section 4.4.3*) were not missed as potential indicators.

4.4.3 Research user group meeting

The Centre's Research User Group (RUG), supported by a RUG Coordinator and Support Assistant, advises and provides feedback on all the Centre's research projects. They contribute to formulating research questions, advise on methods (questionnaire design, recruitment and consent procedures), and contribute to interpreting and disseminating research findings. The objectives of the RUG meeting for this project were:

- To provide older adults with hand pain and functional problems an opportunity to explain their everyday issues related to their hand;
- To ask participants (in small groups) to prioritise indicators that were most applicable to them and their hand condition;
- To identify the indicators/ themes that applied to most/ all of the participants.

Every relevant hand indicator (40 indicators, *Appendix B*) contained in NorStOP was considered as a potential indicator and presented to members of Keele's RUG.

During the RUG meeting, participants were each given the chance to talk about their experiences and history of diagnosis with regard to their hand pain/ problems. The main aim of this was to act as an introduction to the session, but also to get an idea of what was important to each participant. Subsequently the group was split into smaller groups and then asked to look at the 40 indicators provided, and rank them according to what they thought were the most relevant to them and their hand pain/ problems. The following guidance was given for the discussion:

- Filter the indicators to decide which indicators are relevant to you and your hand condition (and exclude those not relevant);
- Rank the remaining indicators based on the impact each indicator has on your hand condition.

After the participants had completed this task, a small group discussion took place with the aim of achieving a consensus of the most important indicators. In addition to this, each group's ranking was recorded, from first to last, so the importance of particular indicators could be investigated at a later date.

In addition to the RUG meeting, a separate private interview was carried out with an elderly family member who received a diagnosis of hand OA approximately 20 years ago, and who generally lives alone. The structure of the conversation was similar to the RUG meeting, where prioritisation of the indicators was discussed.

In order to include as many indicators in the base model development as the participants felt were relevant (without including every indicator), an indicator was included as a potential indicator if at least two groups ranked that indicator in their top 20. It was felt that this approach was sufficient to represent the selections of the participants, whilst filtering out indicators that were not collectively highly ranked.

4.5 Results: Phase I: Identification of indicators important to people with hand OA

4.5.1 Key literature findings

Considering the two PhD theses (Dahaghin, 2005; Kwok, 2013), and additional publications (Kalichman and Hernández-Molina, 2010; Nicholls et al., 2012; Leung et al., 2013), the seven key items associated with hand OA conditions and the progression of the disease that could be directly mapped to available information in the NorStOP were: age; gender; number of OA joints affected; pain intensity; 'pain in both hands'; 'difficulty opening a new jar'; 'difficulty turning taps on'. While there was no item reflecting 'number of OA joints affected' in the NorStOP database, a similar item from the AIMS2 questionnaire ('how often did you have pain in two or more hand joint at the same time during the past month?', with responses 'all days', 'most days', 'some days', 'few days' and 'no days') was deemed comparable and was included as a potential indicator. Age and gender were not included in the list of indicators as these variables would be better suited

to predicting latent state membership or transitions in hand OA state, and did not fit the purpose of identifying phenotypes of hand pain and function. The remaining five indicators were then taken forward to be considered as potential indicators.

4.5.2 RUG findings

The session had eight participants who had some degree of hand pain/ problems, the majority (seven) had symptoms of hand OA; three of these also had received a previous diagnosis of carpal tunnel syndrome, while the remaining participant had been diagnosed with polyarticular gout. All eight participants were aged over 50 years, five were female and three were male. Participants commented on issues around their condition and highlighted what they felt were the biggest aspects that affected them personally.

In the next stage of the session, participants divided themselves into four groups of two, and in their pairs were asked to prioritise the 40 indicators which were printed out as cue cards. The groups generally adhered to guidance for discussion (set out in *4.4.3*), however one group saw the underlying themes within the indicators, and separated their decisions into various themes (which consisted of 'diagnosis indicators', 'functional problems', 'hand feelings and sensations' and 'further specific diagnosis indicators'); the other three groups ranked the ones they felt were most relevant.

As expected, there was variability in the ranking of indicators. All the indicators ranked within the top 20 by any of the RUG sub-groups were compared to get an idea of which indicators were considered important by most or all of the individuals (*Table 4.5.1*). In order to protect confidentiality, the ranking from the additional interview has been anonymously classified as a 'group' in the table.

Table 4.5.1: Distribution of hand indicators by the five groups, indicating number of groups ranking the indicator in top 20 and ranking within each of the groups.

Indicator	Number of groups ranking the indicator within top 20 (out of 5)	Ranking positions within groups
Difficulty opening a new jar	5	1, 6, 6, 7, 11
Difficulty carrying a full pot	5	4, 4, 7, 7, 17
Hand pain in two or more joints	5	5, 8, 11, 18, 20
Pain in both hands	4	1, 1, 7, 17, (NR)
Pain when turning objects in hand	4	4, 6, 12, 13, (NR)
Difficulty turning taps on	3	3, 6, 7, (NR)
Difficulty wringing out a dishcloth	3	5, 6, 9, [22], (NR)
Frequency of medication use for hand symptoms	3	2, 3, 19, (NR)
Pain when gripping objects	3	3, 7, 14, (NR)
Difficulty doing-up buttons	3	4, 9, 16, (NR)
Hobbies with excessive use of hands	3	15, 16, 17, (NR)
Pain when squeezing objects	2	2, 5, [25], (NR)
Write easily with a pen	2	3, 14, [21], (NR)
Burning sensation	2	10, 11, [30], (NR)
Fingers with nodes	2	4, 20, [21], (NR)
Hand stiffness	2	8, 16, [29], (NR)
Hand numbness	2	12, 13, [28], (NR)
Morning hand stiffness	2	15, 17, [27], (NR)

Footnote: Underlined= Indicator not collected at 6 years follow-up, Italics= Not used in base model identification as used to predict membership, NR= Not Ranked by remaining number of groups, []= outside of top 20 ranking.

One of the striking aspects was the number of indicators considered to be of relevance, as over 30 out of the proposed 40 were ranked by at least one group. However, through discussion, these were refined to those agreed as being a concern and/ or having a

reasonable impact on everyday life. Only five indicators ('hand pain in both hands', 'previous 12 month hand pain duration', 'pain when squeezing objects', 'frequency of medication use for hand symptoms' and 'difficulty carrying a full pot'*)* were ranked in the top five of more than one group (*Table 4.5.2*).

Table 4.5.2: The indicators selected in the top five ranking positions by group.

	Group				
Ranking	A	B	C	D	E
1	Job with excessive use	Pain in both hands	Pain in both hands	Feel frustrated about hand problems	Difficulty opening a new jar
2	Difficulty turning a door handle	Taken medication	Easily tie a knot	Previous 12 month hand pain duration	Pain when squeezing objects
3	Difficulty when turning taps on	Duration of hand pain	Pain when gripping objects	Taken medication	Write easily with a pen
4	Pain when turning objects	Fingers with nodes	Difficulty doing-up buttons	Difficulty carrying a full pot	Difficulty carrying a full pot
5	Pain when squeezing objects	Previous hand operation	Difficulty fastening jewellery	Pain in two or more hand joints	Difficulty wringing out a dishcloth

Several functional problems frequently received a high ranking by the groups, namely 'difficulty opening a new jar' and 'difficulty carrying a full pot' (both selected by all five groups), 'pain when turning objects' (selected by four of the groups), 'difficulty turning taps on', 'difficulty wringing out a dishcloth', 'difficulty doing-up buttons' and 'pain when gripping objects' (selected by three of the groups). Other key indicators the participants selected were having 'hand pain in two or more joints' and having 'pain in both hands' (ranked in the top 20 by five and four groups respectively). These indicators were considered as potential indicators when developing the LTA base model (because at least

two groups ranked these indicators in their top 20). Indicators that were ranked in the top 20 by just one group alone have not been included in the table (*Table 4.5.1*).

Similar to age and gender in the highlighted literature (*4.5.1*), a few indicators were removed as potential indictors, as they were better suited to predicting state membership or transitional patterns, and therefore did not fit the aim of identifying phenotypes of hand pain and function (similar to age and gender). The indicators removed were 'frequency of medication use', 'hobbies with excessive use of hands' and 'fingers with nodes', and were not used as indicators in the development of the LTA base model and are indicated as such in *Table 4.5.1* by italics.

4.5.3 Potential indicators for analysis

16 potential indicators remained following the RUG group and key literature process; four indicators identified by both literature and RUG, one by only the literature, and 11 by only the RUG. As there was no indicator for 'pain intensity', e.g. a 0-10 Visual Analogue Scale (VAS), collected at each time point (which was suggested by the literature), it was decided by supervisors and myself that 'pain at rest' would be included in the list of indicators. The main reason for this was it would help to cover an additional area of the construct for hand pain and this linked closest with pain intensity. However, of the 16 indicators, five were not collected in the 6 years survey (indicated in *Table 4.5.1* by an underscore), and therefore are not included in the main analysis (but are included in the pilot analysis, described in *4.6.3*), reducing the number of indicators from 16 to 11.

Therefore, the following 11 indicators were included as potential indicators for the base model of the main analysis (based on relevant factors in the literature and higher priority in the RUG meeting):

- Related to pain:

- Hand pain at rest;
- Hand pain when gripping objects;
- Hand pain when turning objects;
- Hand pain when squeezing objects;
- Hand pain in both hands;

- Related to function:
  - Difficulty turning taps on;
  - Difficulty doing-up buttons;
  - Difficulty opening a new jar;
  - Difficulty carrying a full pot;
  - Difficulty wringing out a dishcloth;
- Related to stiffness:
  - Morning hand stiffness.

4.6 Methods: Phase II: Determining distinct phenotypes of people with hand OA

The objective of the 2nd phase was to carry forward the identified indicators into a modelling procedure to identify phenotypes of hand pain/ problems using the LTA process (described in *Chapter 3*).

4.6.1 Inclusion criteria and definitions

Analysis for the development of the base model used baseline, 3 years and 6 years follow-up data and therefore includes only participants that responded at all three time points. As the investigation is interested in changes in, development of, and resolution of, hand pain/ problems over time, no requirement regarding the presence or severity of hand pain/ problems was made; therefore, participants were included in the analysis even if they had

no hand pain/ problems at all 3 time points. A comparison of baseline characteristics of those who were included in the analysis with those excluded was carried out. By the design of the NorStOP study, if participants did not respond to either the HS or the RPS, they were subsequently excluded from any further follow-up points. For this analysis, participants had to respond to each HS, and each RPS if they were mailed one (and therefore required a RPS) at each of the three time points.

In this analysis, the definition for a participant to be considered to have hand pain/ problems was they indicated as such on the HS questionnaire (hence they would have been sent the RPS), and then to state again they had hand pain/ problems on the RPS. This approach is consistent with previous definitions of 'hand pain/ problems' in the NorStOP study (Hill, 2005). There were no assumptions about participants having hand pain/ problems at any particular time point(s); participants were allowed to have hand pain/ problems at any, all or none of the time points.

Many of the potential indicators were measured using an ordinal scale of 'none', 'mild', 'moderate', 'severe', 'extreme', which were dichotomised between the 'mild' and 'moderate' level, so a '0' score (low) was attributed to 'none' and 'mild', whereas, a '1' score (high) was attributed to 'moderate', 'severe' and 'extreme'; For 'hand pain in both hands', '0' represented hand pain in 'right only' or 'left only' and '1' represented 'both hands'.

For each time point, any participant reporting no hand pain/ problems on either HS or RPS (or both) were given scores for each of the variables of '0', as in each case, a '0' represented none, or mild hand pain/ problems.

It is important to clarify here that participants who reported no hand pain/ problems at the HS, but did in the RPS (hence they must have indicated they had problems in their hip,

knee or foot in the HS), were regarded as having 'none' or 'mild' hand pain/ problems, and their indicators were imputed with a zero. There are two reasons for this: firstly, this was the approach commonly taken when investigating hand pain/ problems in previous NorStOP studies (Hill, 2005; Hill et al., 2007). Secondly, permitting the answers given by participants who filled in the RPS hand indicators but reported no hand problems in the HS would introduce a "comorbidity" bias into the analysis, as these individuals must have had a problem with another condition (hip, knee or foot) to have the opportunity to answer the hand indicators in the RPS. Therefore, using the strict criteria at both stages of the questionnaire process captured the participants with more consistent hand pain/ problems, and also limited bias in individuals with more acute hand conditions. This criterion was enforced at each time point.

Missing values were coded as such, as LTA can account for missing values (EM algorithm, *section 3.9.5*) in the computation process (unless all indicator variables had missing values from each time point investigated and then the observation was excluded from the analysis). Prior to the start of the modelling process, any individuals with missing values for more than half of the potential indicators (so six or more) at any time point were deleted from the analysis, in addition to those who did not respond at each time point.

4.6.2 Modelling process (baseline, 3 years, 6 years)

The details of the modelling process are given below but in general the optimal number of states were identified based on all 11 indicators, then each indicator omitted in turn and the model assessed with that indicator taken out. Then the indicator with the least contribution to the model was removed. This was followed by a check of the number of latent states that were optimum for that model without the omitted indicator, and the process then repeated on the model without that indicator. This process continued until no more terms could be

removed based on model criteria (BIC and entropy), sample size (at least approximately 5% in each latent state) and qualitatively (where the interpretation of each state did not benefit from a removal) (see *section 4.6.4*).

At each stage of the modelling process, the following steps were taken:

1. LTA was performed for all of the variables in the model (so initially 11), to investigate what number of states were optimal, based on BIC, entropy and sample size;
2. For the optimal model, each indicator was removed from the model in turn (so initially there were 11 variants of that model) and model fit (BIC/ entropy, defined in *section 4.6.4.1*/ *4.6.4.2*) compared between these models;
3. Sample size of the states was checked to ensure no state had a small size (approximately <5% of participants in that state);
4. The indicator which best improved the model was removed (biggest reduction in BIC/ increase in entropy);
5. Steps 1 to 4 were repeated until removing further terms provided no further improvement to the model (assessed by looking at the interpretation of the states when taking out a further indicator).

When the last of these stages was reached, a final check was carried out to investigate if including one of the indicators that had already been removed improved the model (via a gain in BIC and an improvement to the interpretation of the states). This stage acted as a 'double-check' to ensure that no indicators that were removed at an early stage should have been returned.

4.6.3 Pilot analysis (baseline and 3 years data only)

A pilot analysis was completed initially with the baseline and 3 years data only for the following reasons. Firstly, by using only the baseline and 3 years data, a larger proportion of the entire population could be used (as the inclusion criteria was adjusted to 'responded to baseline HS/ RPS and 3 years HS/ RPS') as it did not include the participants that were lost to follow-up between 3 and 6 years. Secondly, by only having two time points to analyse, computer processing of the models was much quicker. Therefore, any potential pit-falls in the model development were quicker to arise and the process could be adjusted before the 6 year data (three time points) were included in the analysis. Thirdly, unfortunately, some of the indicators of interest (namely 'hand pain in two or more joints', 'write easily with a pen', 'hand stiffness', 'burning sensation' and 'hand numbness') were not collected at 6 years follow-up. Therefore, undertaking this pilot analysis also allowed the potential for comparison between the 3 years model (with more participants and 16 indicators) and the main 6 years model with the 11 indicators described above.

4.6.4 Criteria for selecting the best model

This next sub-section highlights the techniques used to assess the various models investigated during the model development process. The different approaches/ estimates address different elements of the statistical fit, along with the interpretability of the model. There is no one clear approach for deciding which model is the most suitable, but a trade-off is made between getting the best statistical fit of a model, while being able to provide clear (clinically interpretable) definitions of states that adequately represent the phenotypes.

4.6.4.1 AIC/ BIC

Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) are two of the most common approaches for assessing the fit of Latent Class models (Akaike, 1987; Sclove, 1987). The preferred option of these is the BIC, and is presented by the equation:

$$BIC_{LL_k} = -2LL_k + r\ln(n) \quad \textbf{(4.1)}$$

where LL_k represents the log-likelihood for a model with k groups (clusters/ states), r represents the number of parameters to be estimated and n is the sample size. The BIC estimate is compared between models with different numbers of states and the model with the smallest BIC is optimum. One of the main reasons for the BIC being the more preferred option is that it is an extension of the AIC (explained below) and offers a more conservative estimate (it incurs more of a 'penalty' for a higher number of parameters).

When the number of parameters (r) is large, and sample size is small (state size <5%), then a sample-size adjusted BIC is preferred (Yang, 2006). This equation is the same as (**4.1**) but n is replaced with (n+2)/ 24. The sample size available in this project is relatively large, and the minimum state size is defined as 5%, therefore this measure was not required for the modelling process.

The AIC by Akaike is obtained by the equation:

$$AIC_{LL_k} = -2LL_k + 2r \quad \textbf{(4.2)}$$

where, similar to the BIC, LL_k represents the log-likelihood for a model with k groups and r parameters. Similar to BIC, the model with the smaller AIC estimate when compared with another is optimum. One of the main drawbacks of the AIC is that it is reported to overestimate the number of classes, especially within categorically observed variables

(Yang, 2006; Nylund et al., 2007). Due to the concerns highlighted with AIC from other authors, the BIC was used as the main model criterion for judgement.

4.6.4.2 Entropy

Entropy is an estimate that represents how well the latent classes/ states are defined and are distinct from each other (another interpretation is how much (or how little) there are overlaps between the states). Scores for entropy range from 0 to 1 where a score closer to 1 is optimum, and a lower estimate indicates that distinction between the classes is not well defined. If two states are very similar the entropy of the model would not be very close to 1 as there are clear overlaps in the states. Therefore, entropy is another measure of how well the model fits the data. Similar to BIC, there is no specific target or cut-off estimate for an entropy value; its main purpose is to be used comparatively between similar models for decision-making on number of classes/ inclusion of variables.

A slightly different way of interpreting entropy, according to Collins and Lanza, is that larger values of entropy typically indicate lower classification error (Collins and Lanza, 2010). The authors also provide the following equation (originally proposed by Ramaswamy et al., 1993):

$$entropy = \frac{\sum_{i=1}^{n} \sum_{c=1}^{C} - p_{ic} log\ p_{ic}}{n \log C} \quad \textbf{(4.3)}$$

where p_{ic} represents an individual i's posterior probability of membership in latent class c and n is the sample size. The entropy is the average of all of the individuals' posterior probabilities for their allocated state and values closer to 1 represent better class distinction. Due to the fact that the entropy takes into account all individuals in one estimate representation, it is possible that a few individuals do not assign into states very well, yet overall the separation is good. With this in mind, it is important that entropy is not

used on its own as a model selection tool; entropy should be used in conjunction with other criteria (such as BIC) to support model selection. In relation to this, the average posterior probabilities were checked in the further evaluation of the models described in *Chapter 5* to ensure reasonable classification of the participants in the dataset (*section 5.3*).

4.6.4.3 State sample size

In Latent Class related analyses, where populations are divided into sub-groups, sample size considerations are important. Therefore when assessing the accuracy and fit of models, the sample size of the latent states should be considered to ensure a generalisable model (no very small groups that are uncommon and hence do not represent the population sufficiently). While there is no specific accepted level of class size, generally a minimum of 5% of the total population in each class has been recommended (Nylund et al., 2007). In this dataset, the number of observations was relatively large (over 5,000 observations) and investigating states over 3 time points, so flexibility of this strict cut-off of 5% was allowed in the model development. Therefore each phenotype had to have at least 5% of the population in it on at least one of the time points rather than all three.

4.6.4.4 Interpretation of states/ prevalence of indicator

Another approach to assessing the inclusion/ exclusion of an indicator can be the interpretation of the states surrounding that particular indicator. The main reason for this is that the BIC/ entropy may suggest that the removal of a particular indicator results in an improvement in the model fit; however, keeping the indicator in the model may help the interpretation of the state (and therefore, the label assigned to that sub-group of the population). An indicator that has a consistently low estimated prevalence (item-response probability) or a consistently high prevalence across states, may offer little to the interpretation of the individual states, and therefore may not be a beneficial inclusion in the model. In addition to this, states containing item-response probabilities close to 0.5 do not

help to define the states, as approximately half of the individuals in that state are estimated to be affected by that item, and half are not. Therefore, states with a low number of (or desirably no) indicators with item-response probabilities close to 0.5 were prioritised.

4.6.5 Model estimation

The LTA modelling process was undertaken in Mplus version 7.3 (Muthén and Muthén, 1998-2015) which, for latent method computations, uses the maximum likelihood estimation (where the aim is to find the model solution with the largest log-likelihood value). When LTA (and other latent methodology) is computed in Mplus, specific criteria can be specified to ensure that the model converges on the largest log-likelihood (Muthén and Muthén, 1998-2015; Geiser, 2013). First, random start values are generated by the software (default in Mplus is 10) and used in the model estimation, with Mplus essentially ranking the resulting log-likelihoods in numerical order. The second stage of the process selects the largest log-likelihoods (the number specified by the user) and examines for replication of the log-likelihood. For example, in a model with 500 random start values, and 50 starting value sets, 500 start values will be used to generate log-likelihoods and the 50 largest will be selected. Increasing the number of starting values increases the chance of converging on the largest log-likelihood but also increases the computational time. In addition to this, the user can specify the number of iterations (default is 10) which reflects the number of times Mplus completes the start values/ values selected process. Therefore, increasing the number of iterations can increase the probability of converging on the largest optimum log-likelihood. If the log-likelihood is not selected by at least two sets of starting values, Mplus will issue a warning message as model convergence has not been successfully reached and the user should increase the number of random start values. Using a larger number of random start values and iterations can help to avoid local likelihood maxima (a scenario where an incorrect log-likelihood is deemed the optimum solution, and

can often lead to incorrect parameter estimates). Should the log-likelihood replicate, then this value is used to compute the model estimates.

In the simulations performed in this analysis, 100 random start values, with the 20 largest log-likelihoods selected, and 20 iterations were used. If the log-likelihood did not replicate then the random starts were increased exponentially from 100 to 250, 500, 1000, 5000 and 10,000 (with the number of values selected and iterations increasing from 20 to 25, 50, 100, 500 and 1,000 respectively) until model convergence was achieved. In addition to this, when the final base model was achieved and replicated, that model was re-run with twice the number of random starts to ensure that the model had converged accurately by replicating the same log-likelihood. Should the model not converge after 10,000 random start values, the resulting model information was presented, and this unconverged model was clearly reported. This approach was taken for all Mplus simulations throughout the thesis.

4.7 Results: Phase II: Determining distinct phenotypes of people with hand OA

4.7.1 Characteristics of NorStOP sample

26,705 individuals were identified at the 8 general practices defined in *section 4.3*, of which 26,129 were eligible and sent the HS questionnaire. 18,497 people responded at baseline (70.8%), with 12,847 (69.5%) consenting to further contact. 10,037 (78.1%) participants indicated pain/ problems in one of the four locations of interest and were mailed the RPS. 8,734 (86.8%) returned the RPS questionnaire.

At 3 years, 11,900 were mailed the 3 year HS, and 9,705 (81.6%) responded; 6,622 (68.2%) of these indicated further pain/ problems and were sent the RPS questionnaire, which 5,895 (89.0%) responded.

At 6 years follow-up, 7,637 were mailed the HS, of which 6,423 (84.1%) participants responded. Of these, 4,585 (71.4%) indicated pain/ problems in one of the four locations and were sent the RPS. 4,066 (88.7%) people responded to the RPS of those who were mailed. 5,751 individuals fulfilled inclusion criteria and responded at baseline, 3 years, 6 years, and completed the RPS (if necessary). Of these, a further 134 (2.3%) individuals were removed from the analysis database because they had more than half of their indicators missing at one of the three time points. A flow diagram of the recruitment stages from baseline to 6 years is presented in *Figure 4.7.1*.

The dataset used for the following model selection process was based on the 5,617 participants who met the inclusion criteria. Of these participants, 2,309 reported no hand pain/ problems at any time point, 317 had hand pain/ problems at baseline but not at 3 or 6 years follow-up, 330 had hand pain/ problems at baseline and 3 years but not at 6 years, while 1,372 reported hand pain/ problems at baseline, 3 years and 6 years follow-up.

Comparing the 5,617 participants (BM) included in the base model development with the 12,880 participants (NR) who were not included (due to non-response, refusal or death/ moving practice) revealed a difference in age (mean (Standard Deviation (SD)) BM= 62.64 (8.2) vs. NR= 67.77 (10.6)), indicating the people in the analysis were younger. However, there was little difference in gender between the two groups (female, BM= 54.0% vs. NR= 56.5%). There were differences between the two groups on whether the participants lived alone (BM= 18.9% vs. NR= 27.9%), and their marital and employment status (*Table 4.7.1*).

Figure 4.7.1: Flow diagram of recruitment and retention in the NorStOP, and participants included in this project.

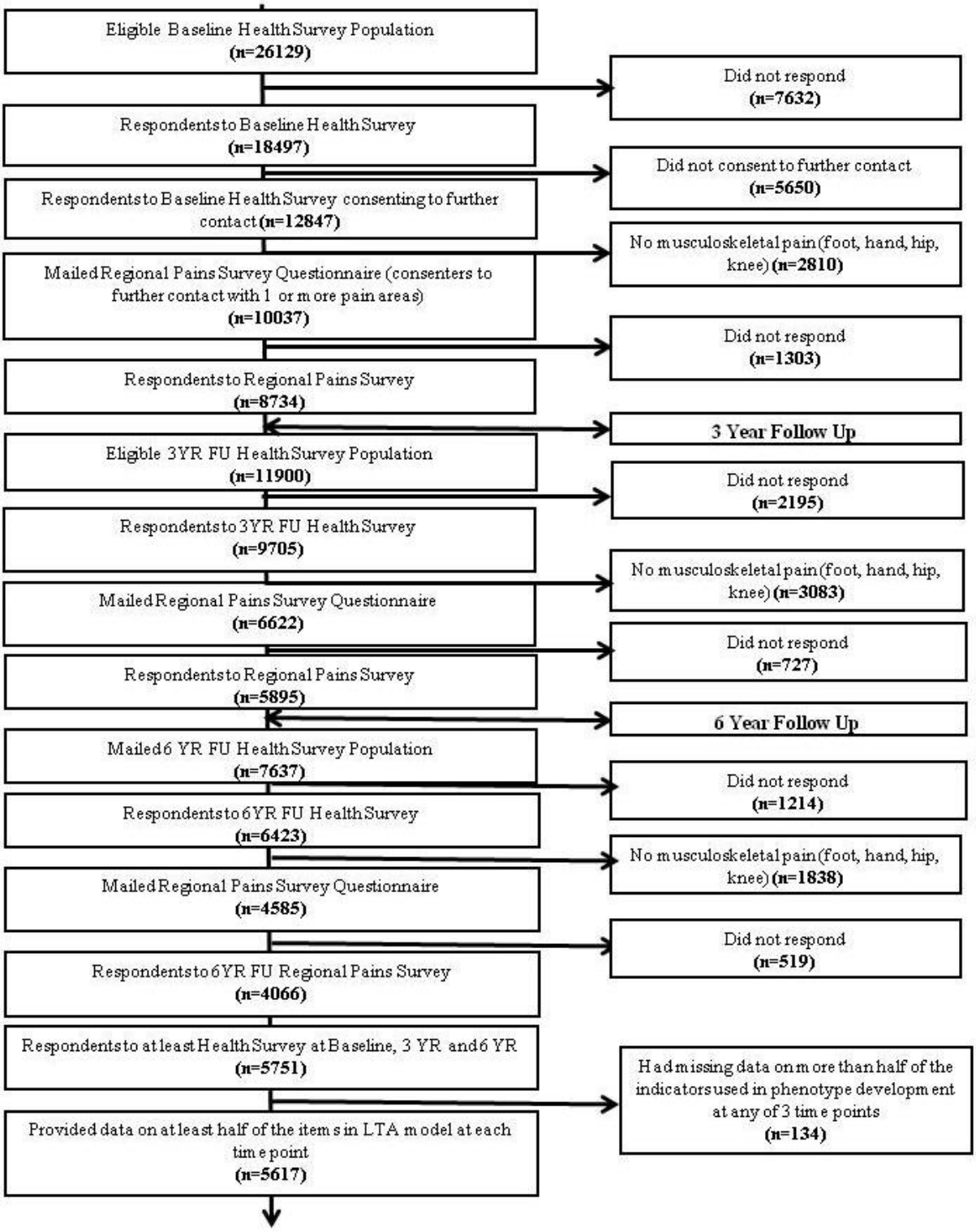


Table 4.7.1: Comparison of baseline characteristics of participants included in the base model development and those lost to follow-up (only completed the baseline HS).

		Base Model participants (BM) n= 5,617	Base Model non-participants (NR) n= 12,880
Age (years, mean (SD))		62.64 (8.16)	67.77 (10.62)
Gender n (%)	Female	3,031 (54.0)	7,280 (56.5)
Lived alone n (%)	Yes	1,020 (18.9)	3,379 (27.9)
Marital status n (%)	Married	4,127 (74.0)	8,071 (63.7)
	Separated	57 (1.0)	103 (0.8)
	Divorced	371 (6.7)	719 (5.7)
	Widowed	652 (11.7)	2,887 (22.8)
	Cohabiting	112 (2.0)	190 (1.5)
	Single	255 (4.6)	699 (5.5)
Employment status n (%)	Employed	2,042 (37.3)	2,819 (22.9)
	Not working due to ill-health	357 (6.5)	965 (7.8)
	Retired	2,549 (46.5)	7,445 (60.4)
	Unemployed	68 (1.2)	107 (0.9)
	Housewife	317 (5.8)	721 (5.9)
	Other	149 (2.7)	258 (2.3)

Footnote: n= number of observations, SD= standard deviation, %= percentage

Table 4.7.2 displays the baseline characteristics of the population used for analysis (n= 5,617) split by being classified as having hand pain/ problems at baseline. It can be seen that those with hand pain/ problems at baseline were significantly more likely to be older (63.0 vs. 62.4 (*p*-value= 0.016)), female (61.6% vs. 49.1%, *p*-value<0.001), to live alone (20.3% vs. 18.0%, *p*-value= 0.037) and to have higher levels of anxiety (mean 7.0 vs. 5.8, *p*-value<0.001) and depression (4.6 vs. 3.5, *p*-value<0.001) symptoms. In addition, participants with baseline hand pain/ problems, were more likely to report sleep disturbances (all *p*-values<0.001) (bottom of *Table 4.7.2*).

Table 4.7.2: Baseline characteristics of those in the base model development analysis, stratified by baseline hand pain/ problems (n (%) unless stated).

Characteristic		Hand pain/ problems at baseline (n= 2,197)	No hand pain/ problems at baseline) (n= 3,420)	*p*-value
Age (years, mean (SD))		62.97 (8.0)	62.43 (8.3)	*p*=0.016
Gender	Female	1,353 (61.6)	1,678 (49.1)	*p*<0.001
Lived alone	Yes	428 (20.3)	592 (18.0)	*p*=0.037
Marital status	Married	1,591 (73.0)	2,536 (74.7)	*p*=0.004
	Separated	17 (0.8)	40 (1.2)	
	Divorced	148 (6.8)	223 (6.6)	
	Widowed	297 (13.6)	355 (10.5)	
	Cohabiting	38 (1.7)	74 (2.2)	
	Single	89 (4.1)	166 (4.9)	
Employment status	Employed	682 (31.9)	1,360 (40.7)	*p*<0.001
	Ill	199 (9.3)	158 (4.7)	
	Retired	1,026 (48.0)	1,523 (45.5)	
	Unemployed	27 (1.3)	41 (1.2)	
	Housewife	139 (6.5)	178 (5.3)	
	Other	63 (3.0)	86 (2.6)	
HADS Anxiety (mean (SD))		6.98 (4.1)	5.81 (3.9)	*p*<0.001
HADS Depression (mean (SD))		4.56 (3.4)	3.47 (3.0)	*p*<0.001
BMI (mean (SD))		27.18 (4.9)	26.5 (4.1)	*p*<0.001
SF-12 general health	Excellent	67 (3.1)	244 (7.2)	*p*<0.001
	Very good	497 (22.8)	1,116 (33.0)	
	Good	929 (42.7)	1,450 (42.8)	
	Fair	571 (26.2)	511 (15.1)	
	Poor	113 (5.2)	64 (1.9)	
Frequency of GP visit	Very often	29 (1.3)	40 (1.2)	*p*<0.001
	Often	362 (16.6)	387 (11.4)	
	Occasionally	1,267 (58.0)	1,733 (50.9)	
	Seldom	357 (16.3)	737 (21.6)	
	Hardly ever	170 (7.8)	511 (15.0)	
Trouble falling asleep	No	786 (36.3)	1,584 (47.2)	*p*<0.001
	Some nights	1,064 (49.2)	1,479 (44.1)	
	Most nights	315 (14.6)	294 (8.8)	
Wake up several times per night	No	264 (12.2)	723 (21.6)	*p*<0.001
	Some nights	1,131 (52.2)	1,836 (54.8)	
	Most nights	771 (35.6)	790 (23.6)	
Trouble staying asleep	No	565 (26.5)	1,314 (39.7)	*p*<0.001
	Some nights	1,066 (50.0)	1,534 (46.3)	
	Most nights	501 (23.5)	465 (14.0)	
Wake up and feel tired	No	655 (30.3)	1,560 (46.6)	*p*<0.001
	Some nights	1,094 (50.6)	1,432 (42.8)	
	Most nights	415 (19.2)	353 (10.6)	

Footnote: n= number of observations; %= percentage; SD= Standard Deviation; HADS= Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983); SF-12= Short Form questionnaire (Ware et al., 1996); GP= General Practitioner.

4.7.2 Model development

4.7.2.1 Pilot analysis - Baseline and 3 years

The modelling process (details of the individual steps are given in the following sub-section for the main analysis of baseline, 3 year and 6 year data) was originally completed using only the baseline and 3 years data (i.e. not the 6 years data). The pilot analysis consisted of 9,705 participants (those who responded to the baseline and 3 year questionnaires), and used the original 16 potential indicator variables, including five that were not collected at 6 years. This pilot analysis produced the same final model as the analysis including the 6 year data (with obviously one less time point and transition matrix) with the same number (and nature) of indicator variables in the final model, and the same number of states/ phenotypes with comparable interpretation of each. The transition matrix was also similar between the pilot and main analysis. As five of the indicator variables were not available for the 6 years analysis, the pilot analysis offered reassurance that the model would not have been affected by the inclusion of these indicator variables. The final model of the 3 year pilot analysis is displayed in *Appendix C*.

4.7.2.2 Main analysis - Reduction of variables

The main point of this sub-section is to highlight how the base model was developed using baseline, 3 year and 6 year data from the indicators highlighted by the RUG and literature. These indicators were reduced down to the main key indicators based on BIC, entropy and qualitative interpretation as explained in *4.6.4*. Table entries in bold in this section represent the optimal model or the next indicator to be removed, depending on the purpose of the table.

At the first stage of modelling, using all 11 indicator variables, a model with five latent states appeared to have a good balance of low BIC (77,612.824), and reasonable entropy

level of 0.910 (*Table 4.7.3*). While the six state model did have a lower BIC, the sample size of some of the states went as low as 3.9% (and had a state that remained below 5% across all time points), therefore a five latent state model (with sample size starting from 5.9%) was chosen for the removal of the first indicator variable.

Table 4.7.3: Model fit parameters with 11 indicator variables for first removal stage.

Number of latent states	BIC	Entropy
2	88,703.406	0.980
3	80,536.693	0.956
4	78,665.192	0.913
5	**77,612.824**	**0.910**
6	77,079.636	0.901

Footnote: BIC= Bayesian Information Criteria; Bold line represents optimum model.

Using a LTA model with five latent states, each of the 11 indicator variables were taken out in turn, the results are displayed in *Table 4.7.4*. The largest reduction in BIC was when the indicator 'pain in both hands' was removed, which also resulted in the biggest improvement in state distinctions (entropy increased to 0.94). Therefore, this was the first indicator removed from the development process. For the next stage of the model development with 10 indicator variables, a five state model was again optimum (*Table 4.7.5*).

Table 4.7.4: Model fit parameters when removing one indicator in turn from the model with 11 indicators and five states.

Removing the term:	BIC	Entropy
Pain in both hands	**66,059.294**	**0.941**
Pain when turning objects	73,392.787	0.901
Pain when squeezing objects	73,472.284	0.899
Pain at rest	71,304.809	0.910
Pain when gripping objects	72,851.986	0.902
Difficulty opening a new jar	72,509.920	0.899
Difficulty carrying a full pot	72,380.162	0.898
Difficulty wringing out a dishcloth	72,668.388	0.904
Difficulty doing-up buttons	73,118.306	0.903
Difficulty turning taps on	73,744.535	0.905
Morning hand stiffness	70,497.258	0.905

Footnote: BIC= Bayesian Information Criteria; Bold line represents optimum model.

Table 4.7.5: Model fit parameters with 10 indicator variables for second removal stage.

Number of latent states	BIC	Entropy
2	75,337.330	0.980
3	68,389.405	0.958
4	67,057.132	0.940
5	**66,059.294**	**0.941**
6	65,946.234	0.895

Footnote: BIC= Bayesian Information Criteria; Bold line represents optimum model.

The next variable to be removed was ‘morning hand stiffness’, with resulting BIC= 58,505.272 and entropy= 0.941 (*Table 4.7.6*), as the removal of this variable led to the biggest reduction in BIC and similar entropy (also the same as the indicator variable ‘pain at rest’).

Table 4.7.6: Model goodness of fit when removing one indicator in turn from the model with 10 indicators and five states.

Removing the term:	BIC	Entropy
Pain when turning objects	61,997.466	0.930
Pain when squeezing objects	61,989.355	0.930
Pain at rest	59,542.061	0.941
Pain when gripping objects	61,398.880	0.934
Difficulty opening a new jar	60,992.594	0.918
Difficulty carrying a full pot	60,999.286	0.919
Difficulty wringing out a dishcloth	61,253.608	0.927
Difficulty doing-up buttons	61,688.125	0.934
Difficulty turning taps on	62,343.371	0.934
Morning hand stiffness	**58,505.272**	**0.941**

Footnote: BIC= Bayesian Information Criteria; Bold line represents optimum model.

Checking the optimum number of latent states again indicated that five was preferred (lowest BIC, lowest class size= 4.5% and all classes >5% for at least one time point). *Table 4.7.7* indicated that the next indicator to be removed was 'pain at rest', with five latent states in the best model, leading to a model with BIC= 51,901.799 and entropy= 0.941.

Table 4.7.8 displays BIC and entropy when removing an indicator from the eight variable model (which has BIC= 51,901.799 and entropy= 0.941). It can be seen that there were some modest reductions in BIC, however, entropy was reduced.

Table 4.7.7: Model goodness of fit when removing one indicator in turn from the model with nine indicators and five states.

Removing the term:	BIC	Entropy
Pain when turning objects	54,506.559	0.931
Pain when squeezing objects	54,464.804	0.931
Pain at rest	**51,901.799**	**0.941**
Pain when gripping objects	53,822.924	0.935
Difficulty opening a new jar	53,472.084	0.927
Difficulty carrying a full pot	53,424.352	0.927
Difficulty wringing out a dishcloth	53,689.913	0.931
Difficulty doing-up buttons	54,059.398	0.933
Difficulty turning taps on	54,821.469	0.931

Footnote: BIC= Bayesian Information Criteria; Bold line represents optimum model.

Table 4.7.8: Model goodness of fit when removing one indicator in turn from the model with eight indicators and five states.

Removing the term:	BIC	Entropy
Pain when turning objects	47,961.139	0.930
Pain when squeezing objects	47,923.288	0.927
Pain when gripping objects	46,993.850	0.933
Difficulty opening a new jar	46,878.505	0.928
Difficulty carrying a full pot	46,836.443	0.925
Difficulty wringing out a dishcloth	47,097.607	0.931
Difficulty doing-up buttons	47,472.253	0.933
Difficulty turning taps on	48,221.168	0.933

Footnote: BIC= Bayesian Information Criteria.

Table 4.7.9 displays a reverse of the process taken so far throughout this section, and represents the BIC/ entropy if one of the indicators previously removed was re-entered into the model. Similarly to before, a lower BIC and higher entropy is optimum. There were two indicators that potentially could benefit from being re-introduced into the model based

on the criteria, namely ‘pain at rest’ and ‘morning hand stiffness’. As there is no specified number of indicators required to be included in the model, the state distinction is a crucial aspect and is an important factor as to whether these indicators were to be re-included or not.

Table 4.7.9: Model goodness of fit when including one indicator to the model with eight indicators and five states, previously removed, in turn.

Including the term:	BIC	Entropy
Pain in both hands	64,056.736	0.904
Pain at rest	51,901.799	0.941
Morning hand stiffness	59,542.060	0.941

Footnote: BIC= Bayesian Information Criteria.

Table 4.7.10 displays the item-response probabilities for the current optimum model, with five latent states and with the indicator ‘pain at rest’ included. The inclusion of this table is to investigate whether ‘pain at rest’ added to the distinction of the classes. The item-response probabilities represent, given that a participant is a member of a particular state, the probability that they responded ‘1’ to each indicator. For example, in *Table 4.7.10*, the top right value of 0.965, represents, for the participants in state five, they had a probability of 0.965 of responding high (moderate, severe, extreme) to the indicator question ‘pain when turning objects’. The item-response probabilities displayed in *Table 4.7.10* (and in the rest of this chapter) are for time 1 only; time 2 and time 3 (3 years and 6 years data) each have their own probabilities estimated. The possibility of restricting the item-response probabilities to be the same at each time point is discussed in *section 5.5*.

Table 4.7.10: Item-response probabilities of model with eight indicators, five states and 'pain at rest'.

	State				
	1	2	3	4	5
Pain when turning objects	0.000	0.575	0.165	0.879	0.965
Pain when squeezing objects	0.002	0.760	0.194	0.941	0.991
Pain at rest	**0.008**	**0.372**	**0.071**	**0.425**	**0.744**
Pain when gripping objects	0.003	0.711	0.148	0.856	0.948
Difficulty opening a jar	0.007	0.163	0.750	0.860	1.000
Difficulty carrying a full pot	0.006	0.057	0.673	0.801	0.982
Difficulty wringing out a dishcloth	0.003	0.141	0.489	0.758	0.988
Difficulty doing-up buttons	0.001	0.027	0.121	0.198	0.872
Difficulty turning taps on	0.000	0.012	0.132	0.106	0.744

Footnote: Bold line highlights the item-response probabilities for 'pain at rest' by state.

The item-response probabilities for 'pain at rest' did not suggest any additional distinction to defining the states when that indicator was included. The probabilities of people reporting 'pain at rest' in states one and three were very low (state three had low probabilities for all the pain items, and state one had low probabilities for all indicators), the probability of reporting pain at rest for those in state five was reasonably high (as are all other indicators in that state). For those in state two there were high probabilities for the indicators reflecting pain during activity, but 'pain at rest' did not have such a high probability. Similarly, in state four high pain during activity indicators appeared prevalent again, yet 'pain at rest' had a low probability in comparison. Finally, the item-response probabilities for 'pain at rest' in states two and four were not well defined (0.37 and 0.43

were reasonably close to 0.5). Therefore, due to this and a lack of qualitative distinction between states for 'pain at rest', and little improvements seen in BIC or entropy, 'pain at rest' was not re-introduced to the model.

In addition to this, a similar investigation was completed exploring re-introducing 'morning hand stiffness' into the model. Similar to 'pain at rest', 'morning hand stiffness' had low item-response probabilities, with the exception of one state (similar to state five in *Table 4.7.10*) where all indicator probabilities were high (and 'morning hand stiffness'= 0.867). The item-response probabilities for 'morning hand stiffness' in all other states were less than 0.43. For similar reasons to 'pain at rest', 'morning hand stiffness' was not reintroduced into the model.

The table below (*Table 4.7.11*) summarises the various removal stages in this sub-section, with the final base model in bold.

Table 4.7.11: Summary table of the stages of model development.

Removal stage	Number of indicators	Indicator removed	Number of states	BIC after removal	Entropy after removal	Smallest state sample size
1	11	-	5	77,612.824	0.910	5.8%
2	10	Pain in both hands	5	66,059.294	0.941	4.6%
3	9	Morning hand stiffness	5	58,505.272	0.941	4.5%
4	**8**	**Pain at rest**	**5**	**51,901.799**	**0.941**	**4.2%**
5	7	Difficulty opening a new jar	5	46,878.505	0.928	3.4%

Footnote: BIC= Bayesian Information Criteria; Bold entry represents the base model.

4.8 The base model

The final model, the 'base model' is displayed in *Table 4.8.1*. Included in this table are the latent state proportions (i.e. the proportion of people assigned to each state, delta's (δ)) and the item-response probabilities (rho's (ρ)) for the five latent states. A more detailed interpretation of the model is provided over the next two chapters where additional elements of model fit are assessed and the characteristics of individuals in each state are presented (along with transition probabilities). However, one additional aspect to highlight at this stage is that the states were assigned labels. The labels were defined in collaboration with a hand researcher outside of the supervisor team (Dr. M Marshall).

The first state (77% of people at baseline) had extremely low probabilities for all item-response probabilities with all $\rho<0.008$. The label 'least affected' was assigned as the participants in this state were least affected by the indicator variables in the base model. This state made up over 70% of the 5,617 people at all time points (although not necessarily the same 70% at each time point) and represented the participants who had little or no hand pain/ problems.

The second state (4.2% of people at baseline) was dictated by the high probability for all the pain indicator variables (where all $\rho>0.64$) whereas the probabilities for the function indicators were moderately small (all $\rho<0.18$). The label 'high pain' refers to the fact that all the pain indicators remaining in the model were high in comparison to the functional items. The third state (5.5% of people at baseline) was titled 'poor gross function' as the functional indicators that could be referred to as representing gross function (e.g. 'difficulty opening a new jar') had relatively high probabilities ($\rho>0.48$). The remaining indicators (e.g. pain and functional items that represent fine motor skills) had comparatively low probabilities ($\rho<0.18$).

Table 4.8.1: Proportion of participants in each state and item-response probability estimates for the final base model with eight indicator variables.

	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Latent state proportions					
Baseline (T1)	0.769	0.042	0.055	0.065	0.070
3 years (T2)	0.717	0.058	0.049	0.095	0.081
6 years (T3)	0.705	0.059	0.046	0.097	0.092
Item-response probabilities					
Pain when turning objects	0.000	0.635	0.147	0.881	0.958
Pain when squeezing objects	0.002	0.827	0.173	0.992	0.989
Pain when gripping objects	0.006	0.710	0.164	0.845	0.938
Difficulty opening a new jar	0.007	0.178	0.742	0.859	0.999
Difficulty carrying a full pot	0.006	0.064	0.661	0.799	0.981
Difficulty wringing out a dishcloth	0.003	0.150	0.483	0.754	0.986
Difficulty doing-up buttons	0.001	0.023	0.118	0.179	0.868
Difficulty turning taps on	0.000	0.013	0.133	0.084	0.884

Footnote: T1= Time 1; T2= Time 2; T3= Time 3.

The fourth state (6.5% of people at baseline) was labelled 'high pain and poor gross function'. This state had a generally similar response pattern as the fifth state, apart from the more detailed 'fine motor' activities which had a lower probability ('difficulty doing-up buttons'= 0.179 and 'difficulty turning taps on'= 0.084). The fifth state (7% of people at baseline) was labelled 'severely affected' as the participants assigned to this state had a high probability of reporting moderate, severe or extreme to all of the indicator indicators, i.e. a response of 1 representing more pain and more functional difficulties.

4.9 Summary

This chapter contained the development of the base model, with the necessary development steps along the way, with a detailed explanation of the study population and data collection of the cohort study used for the analysis in this chapter, and throughout this project. The main analysis in this chapter has defined five distinct hand pain and function phenotypes. The next chapter will assess the accuracy of this model, and explore whether certain restrictions can be placed on this model to aid interpretation.

A limitation of the analysis completed in this chapter is that not all indicators were assessed across all time points, which would have led to the use of 16 rather than 11 indicators being considered at all three time points. However, the pilot analysis completed on the baseline to 3 years data included all of the 16 indicators selected from the RUG and literature, and the indicators missing from the 6 years data did not appear in the pilot model. Therefore, this does provide some reassurance that these indicators would not have been included in the 6 years states.

An additional limitation is that while the software can incorporate missing data into the modelling process, some observations with missing data were included. Individuals with missing indicator responses were designated into states similar to individuals with similar indicator responses, for those indicators without missing values. Therefore, this may have resulted in some individuals not being designated to their optimum state. This potential issue is explored in more detail in the next chapter. However, the main approach taken to try to reduce the effect of missing data was to exclude any individual with missing data on more than half of the indicators at any time point before the model development began. This approach was taken to prevent individuals being classified into states based on a low number of indicator responses (and therefore increasing the possibility of incorrect state

designation). However, individuals with missing data on less than half of the indicators were then permitted for analysis and estimated using the EM algorithm (described in *section 3.9.5*). In conjunction to this, none of the individuals in the analysis dataset (5,617) had missing data on five or more indicators at any time point, while only 283 (5% of 5,617) had any missing data.

Also, the decision was made to dichotomise the indicators into a binary form (often from five responses). The main reasons for this were to speed up the computation of the models and aid interpretation of the results. There are several issues with taking this approach. Firstly, it may be the case that splitting the five levels into a binary representation is too simplified, and there may be participants with more complex patterns that a binary indicator cannot represent. Secondly, it was decided to dichotomise between the 'none and mild' and 'moderate, severe and extreme' levels of the indicators. It may be that this is not the optimum point to dichotomise the indicator, and including 'moderate' with 'none and mild' might have been more appropriate. A sensitivity analysis is performed at the end of the thesis to explore the impact of adjusting the dichotomisation (*Chapter 10*).

The next chapter further assesses the states identified in the base model and whether state interpretations remain the same at each time point.

Chapter 5: **Assessment of the base model**

5.1 Introduction and objectives

The purpose of this chapter is to assess whether the states identified in the base model developed in *Chapter 4* appear to give a reasonable representation of the observed hand pain/ problem characteristics of the older population. In addition, another key focus of this chapter is to assess whether restricting the item-response probabilities to be the same at each time point is reasonable, thus ensuring phenotype definition remains stable at each time point.

The specific objectives of this chapter were to:

i. Assess the potential influence of non-response on the identified states by determining whether respondents to the 6 year analysis were considered comparable to those who failed to respond at 6 years;
ii. Assess whether individuals were clearly allocated into a latent state at each time point by assessment of the average posterior probabilities;
iii. Assess if the distribution of observed indicator responses of people allocated to each state were similar to that state's estimated item-response probabilities at baseline;
iv. Investigate if it was reasonable to restrict the item-response probabilities to be the same at each time point, and therefore have the same latent state interpretation at each time point;
v. Examine the 'least affected' state in more detail to ensure this was a valid and distinct state as this included both people reporting no hand

pain/ problems and those reporting hand pain/ problems but with limited impact.

5.2 Non-response

As mentioned in *4.6.3*, a pilot analysis was originally completed on the available 3 years data, before the additional 6 years data was added. The pilot 3 years data analysis revealed the same number of states with similar state definitions as the 6 years data. It is possible to explore the states of individuals from the pilot analysis that were lost to follow-up at 6 years which permits a comparison of state proportions between those in the 6 years data analysis and those lost to follow-up regarding the potential for non-responder bias. This bias is likely when individuals who have not continued in the study (through non-response) are different in some way to those who completed the 6 year questionnaire. However, considering the similarities between the latent states developed in the pilot analysis reported in *section 4.7.2.1*, and base model (*4.8*), we may presume that these models should be similar.

Table 5.2.1 compares the 4,088 that were lost to follow-up at 6 years (the 'Non-responders at 6 years'), to the 5,617 individuals that were included in the 6 years analysis (the 'Responders at 6 years') on their latent state membership in the pilot analysis. One of the main findings to highlight is that there was a higher proportion of individuals in the 'severely affected' state amongst the 6 years non-responders. 11.4% of non-responders were in the 'severely affected' state at baseline and 13.7% at 3 years; for responders at 6 years, there were 6.4% and 8.6% in this state, respectively. Therefore, it does appear that a higher proportion of individuals in the worst state were lost to follow-up.

Table 5.2.1: Comparison of latent states in the pilot analysis between those lost to follow-up at 6 years (n= 4,088) and those in the full 6 year analysis population (n= 5,617).

	Non-responders at 6 years (n= 4,088)		Responders at 6 years (n= 5,617)	
	Time 1	Time 2	Time 1	Time 2
LA	2,884 (70.5%)	2,840 (69.5%)	4,275 (76.1%)	4,003 (71.3%)
HP	157 (3.8%)	132 (3.2%)	246 (4.4%)	328 (5.8%)
PGF	279 (6.8%)	122 (3.0%)	341 (6.1%)	250 (4.5%)
HPPGF	301 (7.4%)	432 (10.6%)	397 (7.1%)	554 (9.9%)
SA	467 (11.4%)	562 (13.7%)	358 (6.4%)	482 (8.6%)

Footnote: LA= 'least affected'; HP= 'high pain'; PGF= 'poor gross function'; HPPGF= 'high pain and poor gross function'; SA= 'severely affected'.

5.3 Average posterior probabilities

A common approach in latent class methods to ensure that participants have been designated into the correct latent state at each time point is through examining the average posterior probabilities (APP) of state membership. As mentioned in *section 3.3*, individuals are classified into each state based on their largest posterior probability. If individuals have similar probabilities for more than one state, this suggests poor discrimination between states. For example, in a five state model, an individual could have a 0.5 probability of belonging in state one and 0.5 of state two (and 0.00 for state three, four and five). This would indicate that, for this individual at least, the model is not sufficient to classify them successfully as they could be allocated to state one or two. APPs of greater than 0.7 for individuals allocated in each state are considered to indicate distinct state designation (Clark et al., 2006).

Table 5.3.1 shows the average probability of individuals being designated into each state compared against the state they were actually classified into. For example, for individuals classified into the 'least affected' state at baseline (time 1), they, on average, had a very

high (0.99) posterior probability of being classified in that 'least affected' state at baseline and minimal probabilities (<0.01) of being classified into the other states.

Table 5.3.1: Assessment of the average posterior probabilities for each latent state against actual state classification for each time point (n= 5,617).

State classified in at:	Average posterior probabilities for each state				
Time 1	LA	HP	PGF	HPPGF	SA
LA	**0.99**	0.00	0.00	0.00	0.00
HP	0.02	**0.89**	0.04	0.05	0.00
PGF	0.04	0.04	**0.87**	0.05	0.00
HPPGF	0.00	0.05	0.05	**0.86**	0.05
SA	0.00	0.00	0.01	0.07	**0.92**
Time 2	LA	HP	PGF	HPPGF	SA
LA	**0.99**	0.00	0.01	0.00	0.00
HP	0.04	**0.86**	0.03	0.06	0.00
PGF	0.04	0.05	**0.88**	0.03	0.00
HPPGF	0.00	0.05	0.02	**0.87**	0.06
SA	0.00	0.00	0.00	0.07	**0.93**
Time 3	LA	HP	PGF	HPPGF	SA
LA	**0.99**	0.00	0.01	0.00	0.00
HP	0.01	**0.89**	0.02	0.07	0.00
PGF	0.03	0.04	**0.89**	0.04	0.00
HPPGF	0.00	0.05	0.02	**0.85**	*0.08*
SA	0.00	0.00	0.00	0.05	**0.94**

Footnote: Bold entries represent APP for optimum state classification; LA= 'least affected'; HP= 'high pain'; PGF= 'poor gross function'; HPPGF= 'high pain and poor gross function'; SA= 'severely affected'.

All of the diagonal entries (in bold) in *Table 5.3.1* contain the APP of individuals being in the state they were classified into. As desired, these probabilities were all high (>0.70) for each state at each time point, indicating that the states appeared to be distinct and individuals were clearly allocated to their optimum state. Consequently, the probabilities for being in a different state from the one allocated were generally small. The highest was 0.08 (in *Italics* in *Table 5.3.1*), suggesting that there may be a small number of people that were classified in the 'high pain and poor gross function' that could have been designated to the 'severely affected' state. However, we know from the exploration of the item-response probabilities that these two states had some similarities, so this slight overlap is not unexpected.

5.4 Indicator response associated with the latent states

The base model defined in *Chapter 4* reduced the potential variables down to eight indicators, with five latent states over the three time points (baseline, 3 and 6 years). This section explores the distribution of responses to the indicators in the base model, to assess the face validity of the model, to describe the frequency of pain and function limitations in each latent state and to compare the observed to expected probabilities of response to the indicators.

Table 5.4.1 details the observed and estimated distribution of responses to the indicators, for each of the latent states. The purpose of this table is to assess for each latent state if the item-response probabilities were similar to the observed proportions of people reporting a high score ('moderate', 'severe' or 'extreme' response) in people allocated to that state. The results (*Table 5.4.1*) show the proportion of people in each state reporting high on an indicator were similar to the item-response probabilities for that state.

For example, amongst the participants allocated to the forth latent state, 'high pain and poor gross function', at least 73% of respondents reported 'moderate', 'severe' or 'extreme' pain or function limitation for the first six pain and gross function indicators, and their corresponding item-response probabilities were all greater than 0.75. Due to the similarities between the proportions in *Table 5.4.1*, and the probabilities displayed underneath each, it could be assumed that the model provided a good representation of actual states of hand pain/ problems. The item-response probabilities also provided evidence that the states were different to each other (i.e. no two states had very similar item-response probabilities across indicators), and there were no two obvious states that would benefit from being merged. Similar results were seen when comparing proportions

of people responding positively to each indicator and item-response probabilities for follow-up measurements at 3 and 6 years (not displayed).

Table 5.4.1: Frequency of 'moderate', 'severe' or 'extreme' responses (with percentages) to each indicator, paired with item-response probabilities, by each latent state at baseline, (n (%) for indicators).

Indicator in base model	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Observations in state (n)	4337	214	305	356	405
Pain when turning objects	0 (0.0)	141 (65.9)	39 (13.0)	315 (89.7)	379 (96.2)
Item-response probability	*0.000*	*0.635*	*0.147*	*0.881*	*0.958*
Pain when squeezing objects	14 (0.3)	185 (86.5)	50 (16.6)	332 (94.3)	389 (99.0)
Item-response probability	*0.002*	*0.827*	*0.173*	*0.992*	*0.989*
Pain when gripping objects	30 (0.7)	156 (72.9)	40 (13.4)	304 (86.6)	369 (94.1)
Item-response probability	*0.006*	*0.710*	*0.164*	*0.845*	*0.938*
Difficulty opening a new jar	30 (0.7)	30 (14.1)	232 (76.6)	304 (85.9)	403 (100.0)
Item-response probability	*0.007*	*0.178*	*0.742*	*0.859*	*0.999*
Difficulty carrying a full pot	24 (0.6)	10 (4.7)	205 (67.9)	282 (79.7)	396 (98.0)
Item-response probability	*0.006*	*0.064*	*0.661*	*0.799*	*0.981*
Difficulty wringing out a dishcloth	12 (0.3)	31 (14.5)	152 (50.2)	261 (73.5)	399 (98.8)
Item-response probability	*0.003*	*0.150*	*0.483*	*0.754*	*0.986*
Difficulty doing-up buttons	7 (0.2)	5 (2.3)	35 (11.5)	59 (16.6)	348 (85.9)
Item-response probability	*0.001*	*0.023*	*0.118*	*0.179*	*0.868*
Difficulty turning taps on	0 (0.0)	2 (0.9)	42 (13.8)	14 (3.9)	362 (89.8)
Item-response probability	*0.000*	*0.013*	*0.133*	*0.084*	*0.884*

Footnote: %'s are of available data within each latent state, n= number of observation; Entries in Italics are item-response probabilities for same indicator.

5.5 Restricting item-response probabilities

An important aspect of LTA concerns the stability of the item-response probabilities (rho (ρ)) over the time periods investigated. Many studies enforce restrictions to ensure identical definition of states over time in order to facilitate a more straightforward interpretation of transitions between states over time. If item-response probabilities cannot be restricted over time, this will complicate interpretation because stability in one latent state between two time points (staying in the most comparable state), does not necessarily mean the condition of that individual has remained stable. However, it is important to check that this restriction can be justified, statistically and qualitatively.

There are few accepted approaches to checking if restricting item-response probabilities to be the same over time is justified (Collins and Lanza, 2010). The two models considered were the model with no restrictions (the base model displayed in *Table 4.8.1)* which for this sub-section is denoted as the 'null model', and the model where the item-response probabilities were restricted to be the same over the three time points, denoted as 'restricted model'. The first check is through comparing the BIC between the two models (explained in *4.6.4.1*), and similar to its usage throughout *Chapter 4*, the model with a lower BIC is preferred.

A second method for assessing if the restricted model is suitable is through a LRT. As these models were essentially the same except one was more restrictive that the other, they could be considered as nested models. The information on log-likelihood and degrees of freedom can be used to test for a statistical difference between the two nested models using the following LRT equation:

$$Difference = -2 \ln \left(\frac{likelihood\ for\ 'null\ model'}{likelihood\ for\ 'restricted\ model'} \right) \quad \textbf{(5.1)}$$

which can be written as:

$$= -2\ln(likelihood\ for\ null\ model) + 2\ln(likelihood\ for\ restricted\ model) \quad (\mathbf{5.2})$$

To note, the likelihood ratio statistic (sometimes referred to as G^2), used to assess model fit in latent class models, is defined as:

$$G^2 = 2\sum_{w=1}^{W} f_w log\left(\frac{f_w}{\hat{f}_w}\right) \quad (\mathbf{5.3})$$

where W represents the number of parameters estimated, f_w represents the observed frequency of cell w, and $\hat{f}_w$ represents the expected frequency of cell w according to the fitted model.

The difference between the two nested models using the equation above is then tested on a Chi-squared (χ^2) distribution (which has been deemed to reflect the G^2 distribution well, Collins and Lanza, 2010), on (twice) the difference in degrees of freedom (*df*) between the two models, simply calculated by the equation:

$$2 \times Difference\ in\ df = df\ in\ 'null\ model' - df\ in\ 'restricted\ model' \quad (\mathbf{5.4})$$

If the resulting *p*-value is significant, then it is deemed that restricting the item-response probabilities is of significant harm to the model, and the test advises against it. However, if the test is non-significant (p>0.05), the additional step of restricting the probabilities does not significantly harm the model and the parameter restrictions can be enforced. It is important to note that when a large number of parameters (*df*) are involved, the distribution of G^2 may not be approximated well by the χ^2. Therefore, it is extremely useful to assess the BIC in addition to the LRT (Collins and Lanza, 2010).

A third check is to visually compare the estimated item-response probabilities for each indicator at each time point in the null model by state (item-response probabilities in the restricted model will be identical at each time point by design), and explore whether the estimates are similar at each time point. A way of checking whether this holds is to query whether each latent state can plausibly be given the same label definition at each time point considered; so does the state label justifiably remain the same over time?

As there is no one declared 'gold standard' to determine whether restricting the probabilities can be presumed, if the majority (2 out of 3) of these tests implied that restricting the item-response probabilities was a reasonable assumption, this was enforced due to the benefit of model interpretation and in further model investigations. All of these approaches are discussed and recommended in Collins and Lanza, 2010.

The BIC of the model with the restricted probabilities was somewhat smaller than the null model with no restrictions (51,349 vs. 51,902) (*Table 5.5.1*); therefore, the restricted model was preferred in terms of BIC assessment.

The difference in log-likelihoods between the two models was tested. The likelihood difference was

$$= -2(-25{,}311.754) + 2(-25{,}242.948)$$

$$= 50{,}623.508 - 50{,}485.896$$

$$= 137.612$$

with

$$2(164 - 84) = 160\ degrees\ of\ freedom$$

On a chi-squared distribution, a value of 137.612 with 160 degrees of freedom produces a *p*-value= 0.899, which implied that including the restrictions on the model, for the LRT, was not of significant harm to the model.

Table 5.5.1: Summary of information for both null and restricted model.

	Null model	Restricted model
BIC	51,902	51,349
Entropy	0.941	0.940
Log-likelihood	-25,242.948	-25,311.754
Degrees of freedom	164	84

Footnote: BIC= Bayesian Information Criteria.

Table 5.5.2 displays the item-response probabilities for each indicator in the final model at all 3 time points, for each latent state in the null model. It can be seen that, with the exception of the 'poor gross function' state, the probabilities for the pain indicators were, in general, similar or slightly increased at each time point. The indicators that related to gross function appeared to remain the most constant estimates across the time period, with some small increases in the probabilities of poorer gross function in the 'high pain' state (implying that the gross function of those classified in this latent state were gradually deteriorating).

For the functional indicators that were related to fine motor skills (doing-up buttons, turning taps on), in the 'severely affected' and 'high pain and poor gross function', the probability estimates slightly increased over the 3 time points, demonstrating that these skills were worsening in the individuals classified in those states. Despite (subtle) changes in these estimates over the three time points, the same state labels and definitions could be applied at each time point. Therefore, the final assessment of restricting item-response probabilities was also satisfied in favour of restriction.

Considering all information for the two models (that is summarised in *Table 5.5.1*), item-response probability restrictions were enforced as this model produced a lower BIC, a non-significant LRT test and little qualitative differences over the three time points.

Table 5.5.2: Item-response probabilities for each indicator at all three time points split by latent state in the null model.

Indicator		Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Pain when turning objects	Time 1	0.000	0.635	0.147	0.881	0.958
	Time 2	0.001	0.737	0.112	0.925	0.971
	Time 3	0.002	0.805	0.106	0.921	0.992
Pain when squeezing objects	Time 1	0.002	0.827	0.173	0.943	0.989
	Time 2	0.003	0.760	0.110	0.963	0.981
	Time 3	0.005	0.880	0.167	0.963	0.994
Pain when gripping objects	Time 1	0.006	0.710	0.164	0.845	0.938
	Time 2	0.007	0.747	0.144	0.853	0.970
	Time 3	0.006	0.821	0.124	0.888	0.996
Difficulty opening a new jar	Time 1	0.007	0.178	0.742	0.859	0.999
	Time 2	0.002	0.232	0.661	0.878	0.998
	Time 3	0.006	0.264	0.760	0.940	1.000
Difficulty carrying a full pot	Time 1	0.006	0.064	0.661	0.799	0.981
	Time 2	0.003	0.075	0.586	0.809	0.998
	Time 3	0.004	0.123	0.614	0.852	0.993
Difficulty wringing out a dishcloth	Time 1	0.003	0.150	0.483	0.754	0.986
	Time 2	0.002	0.188	0.355	0.793	0.991
	Time 3	0.000	0.205	0.472	0.795	0.988
Difficulty doing-up buttons	Time 1	0.001	0.023	0.118	0.179	0.868
	Time 2	0.001	0.050	0.157	0.225	0.914
	Time 3	0.001	0.038	0.239	0.300	0.940
Difficulty turning	Time 1	0.000	0.013	0.133	0.084	0.884

taps on	Time 2	0.000	0.020	0.055	0.143	0.871
	Time 3	0.000	0.012	0.081	0.210	0.918

Therefore, from this point forward the 'base' model used is the restricted model, with item-response probabilities identical at each time point, so states can be given the same interpretation at each time point and ease the overall interpretation of the model.

5.6 Least affected state

Due to the inclusion criteria, individuals who did not report hand pain/ problems in the HS at any time point were still included in the analysis, and were automatically recorded as having no problem for all their indicator questions for that time point. The 'least affected' state could have therefore included these individuals who did not indicate hand pain/ problems but also included some reporting hand pain/ problems but have a limited impact from these.

At baseline, 4,338 individuals were classified as being in the 'least affected' state. Of this number, 918 (21.2%) reported hand pain/ problems in the baseline HS and RPS but reported 'none' or 'mild' problems on the indicator questions. Of the remaining 3,420 individuals, 350 (8.1%) reported hand pain/ problems in the HS but not in the RPS, 237 (5.5% of state population) reported hand pain/ problems in the RPS but not in the HS (they must have indicated problems in another location to qualify to receive the RPS), and the remaining 2,833 (65.3%) reported no hand pain/ problems in either HS or RPS. Therefore, over three quarters (78.8%) of the 'least affected' state at baseline were made up of individuals who did not report hand pain/ problems at baseline in at least one of the HS and RPS, and subsequently were automatically given a 'none'/ 'mild' ('0') entry for responses to the indicators.

Of the 4,068 individuals in 'least affected' state at time 2 (3 years), 897 (22.1%) reported hand pain/ problems at both questionnaire stages, 2,625 (64.5%) at neither stage, 276 (6.8%) hand pain/ problems at the HS only, and 270 (6.6%) at the RPS only. So at 3 years, 77.9% of the state was configured of individuals who did not report hand pain/ problems in at least one of the HS and RPS at time 2.

At the third time point (6 years), of the 3,966 in the 'least affected' state, 665 (16.8%) reported hand pain/ problems in both HS and RPS. 2,748 (69.3%) did not report hand pain/ problems in either HS or RPS, 389 (9.8%) reported hand pain/ problems in the HS only, and 164 (4.1%) reported pain/ problems in the RPS only. As such, at this time point, 83.2% of the state was made up of individuals who did not report hand pain/ problems in at least one of the HS and RPS at 6 years follow-up.

There did appear to be a larger proportion of individuals at 6 years who reported no hand pain/ problems in both the HS and the RPS compared to at baseline and 3 years. One potential explanation for this discrepancy could be that individuals who reported mild hand pain/ problems at baseline and 3 years but were allocated to the least affected state, transitioned into one of the problematic states at 6 years.

In conclusion, the majority of individuals classified as 'least affected' reported no hand pain/ problems in the HS and/ or RPS in each respective time point. Of the remainder who did report pain or problems, most indicated 'none' or 'mild' symptoms to the indicators included in the base model for each respective time point (*Table 5.4.1*). Therefore, this state did appear to be a valid and distinct state which contained individuals with no or mild symptoms of hand pain/ problems.

5.7 Summary

This chapter has assessed the base model developed in *Chapter 4* ensuring individuals were generally designated into their optimum latent state. In addition to this, the assumption of restricting the item-response probabilities over time was assessed, with the conclusion that the interpretation of the latent states could be considered the same at each time point.

Comparing the distribution of the observed responses with item-response probabilities for each indicator by each latent state offered further assurance that the model and the states represented the patterns of hand pain/ problems observed amongst the participants. Similar to this, exploring the average posterior probabilities gave confidence that individuals had been clearly classified into their optimum state at each time point, with high probabilities of being designated into their allocated state, and low probabilities of being in a different state. Therefore, these two investigations supported the premise that the latent states matched the hand pain/ problems profiles of participants, were distinct, and had correctly classified individuals into their most likely states.

One identified potential limitation was that some of the individuals with more severe hand pain/ problems may have been lost to follow-up between 3 and 6 years. A consequence of this could be attrition bias, in that some of the individuals with more advanced or severe hand pain/ problems were not represented in the 6 years analysis. Another potential limitation was the discovery that there were a slightly higher proportion of individuals in the least affected state who reported no hand pain/ problems at 6 years, compared to baseline and 3 years (*section 5.6*). The 'least affected' state (at each time point) was a combination of individuals that represented various response characteristics to the HS and RPS over the time points, but essentially were individuals stating no/ little hand pain/

problems. These issues could reflect a loss to follow-up of individuals with more severe problems, which could potentially result in higher rates of transition from the more severe states to 'least affected' over time. This potential pattern is explored further in *Chapter 6*.

The next chapter presents further details of the restricted LTA model, initially by describing the demographic characteristics of individuals in each hand phenotype. Next, the transition probabilities between states from baseline to 3 years, and 3 years to 6 years are presented. The remainder of the chapter includes two LTA extensions highlighted in the literature review: transition probabilities stratified by participant characteristics (such as age or gender), and including covariates when computing the latent states and subsequently assessing potential improvements in model fit.

Chapter 6: **Further interpretation of the base model**

6.1 Introduction and objectives

This chapter follows directly on from *Chapter 5*, by exploring the restricted base model further. The main focus of this chapter was to explore how the baseline demographics and lifestyle factors of respondents varied across the latent states developed over the last two chapters, determine patterns of transition over the three time points, and investigate if there were any differences in transitional patterns according to key demographic information. Further to this, the last section of this chapter explores if including a covariate into the modelling process resulted in an improved model fit which could indicate important characteristic differences between the phenotypes.

The specific objectives of the chapter were to:

i. Assess demographic and general health characteristics associated with latent state membership at baseline;
ii. Identify the main transition patterns over 6 years;
iii. Examine if there were any significant differences in transitional patterns by gender, living status, age and widespread pain;
iv. Explore the inclusion of confounders into the latent state development model to assess model fit.

6.2 Demographic and general health characteristics of latent states at baseline

6.2.1 Methods

To begin exploration into the factors that were associated with membership within each latent state at baseline, the first stage was to identify the potentially important sociodemographic and general health factors collected in the baseline HS questionnaire. In addition to demographic variables (age, gender, marital and employment status), the effect of living status (whether the individual lived alone or not) was examined as RUG members suggested that those who lived alone found it more difficult to cope with their hand condition (*Chapter 4*). Further to this, presence of widespread pain based on the American College of Rheumatology (ACR) criteria (Wolfe et al., 1990) was investigated in more detail as it was theorised (with supervision team) that those with widespread pain were potentially more likely to be affected by their hand condition; this has been suggested in previous literature (Grotle et al., 2008a; Nicholls et al., 2012).

Poor mental health has been shown to be associated with painful musculoskeletal conditions (Bair et al., 2003; Arnow et al., 2006; He et al., 2008; Arola et al., 2010; Kroenke et al., 2011), and therefore measures for anxiety and depression symptoms were included in this analysis using the HADS index (Zigmond and Snaith, 1983). The HADS index contains 14 indicators, seven for each of anxiety and of depression, and provides a score for each condition on a scale of 0-21. A higher score indicates the respondent is more likely to have anxiety/ depression, while those in the mid-range (scores ranging between eight and 10) are regarded as having possible anxiety/ depression (Zigmond and Snaith, 1983). The HADS index has been found to have good psychometric properties in primary care samples (Pallant et al., 2005; Bunevicius et al., 2007; Cameron et al., 2008).

There is inconsistent evidence of an association between BMI and the development of hand OA (Carmen et al., 1994; Haara et al., 2003; Kalichman and Kobyliansky, 2007; Bernard et al., 2010; Yusef et al., 2010; Ghosh et al., 2014; Prieto-Alhambra et al., 2014), therefore BMI was included in the analysis to further investigate its association with hand OA.

Chapter 8 investigates consultation frequency for hand problems and how they related to the distinct latent states. However one question within the HS ('how often do you visit the doctor (GP) for yourself?') was included in this descriptive analysis. This item represented a self-reported subjective assessment of how frequently an individual visited their GP (for anything), and allowed an assessment of the association of general health care use with the latent states. Self-reported health was assessed using an item from the SF-12 ('In general would you say your health is: 'excellent', 'very good', 'good', 'fair' or 'poor'?'), which was included here to assess how individuals with various stages of hand pain/ problems regarded their overall health (Ware et al., 1996). Finally, various aspects of sleep measured using the Jenkins scale (Jenkins et al., 1988), were assessed because during the RUG meeting, some members highlighted they often found it difficult to fall asleep due to their hand pain.

All of the baseline investigations by latent state are presented as means with standard deviations for continuous measures, and number with percentages for categorical measures. Significance tests were carried out to determine unadjusted associations between baseline state membership and each of the baseline variables (χ^2 test for categorical variables, one way ANOVA (Analysis of Variance) for continuous variables).

6.2.2 Results

Table 6.2.1 displays the demographic information of the participants in the base model, stratified by latent state at baseline. A higher proportion of females and older participants were in the more detrimental states ('severely affected', 'high pain and poor gross function' and 'poor gross function'). Participants in these states were also more likely to report living alone, visiting the GP for any condition, have higher levels of anxiety and depression symptoms, report widespread pain and sleep problems. In terms of self-reported general health, there was a clear trend between increased severity of hand pain/ problems and reduced self-rated health.

The participants in the 'least affected' and 'high pain' states at baseline were more likely to be married and in employment (potentially due to them being younger), and have lower BMI's. All of the factors assessed were significantly different between latent states at baseline.

Table 6.2.1: Baseline demographic and general health characteristics by baseline latent state (n (%) unless stated).

Baseline factor		LA	HP	PGF	HPPGF	SA	*p*-value
Observations (n)		4338	224	307	394	354	
Age (mean (SD)) (n= 5,617)		62.4 (8.2)	62.1 (7.4)	64.2 (8.1)	63.3 (7.9)	64.4 (8.1)	<0.001
Gender (n= 5,617)	Male	2171 (50.1)	128 (57.1)	66 (21.5)	128 (32.5)	93 (26.3)	<0.001
Lived alone (n= 5,408)		752 (18.0)	29 (13.6)	60 (20.1)	85 (22.1)	94 (28.3)	<0.001
Marital status (n= 5,574)	Married	3238 (75.2)	178 (79.8)	218 (71.2)	279 (71.2)	214 (61.3)	<0.001
	Separated	45 (1.1)	3 (1.4)	3 (1.0)	3 (0.8)	3 (0.9)	
	Divorced	272 (6.3)	10 (4.5)	24 (7.8)	29 (7.4)	36 (10.3)	
	Widowed	457 (10.6)	17 (7.6)	41 (13.4)	59 (15.1)	78 (22.4)	
	Cohabiting	85 (2.0)	4 (1.8)	8 (2.6)	8 (2.1)	7 (2.0)	
	Single	207 (4.8)	11 (4.9)	12 (3.9)	14 (3.6)	11 (3.2)	
Employ-ment status (n= 5,482)	Employed	1740 (41.0)	92 (41.6)	78 (26.1)	88 (23.1)	44 (13.0)	<0.001
	Ill	184 (4.3)	13 (5.9)	28 (9.4)	55 (14.4)	77 (22.8)	
	Retired	1915 (45.1)	96 (43.4)	163 (54.5)	188 (49.3)	187 (55.3)	
	Unemployed	54 (1.3)	6 (2.7)	1 (0.3)	7 (1.8)	0 (0)	
	Housewife	233 (5.5)	5 (2.3)	25 (8.4)	33 (8.7)	21 (6.2)	
	Other	117 (2.8)	9 (4.1)	4 (1.3)	10 (2.6)	9 (2.7)	
HADS Anxiety (mean (SD)) (n= 5,527)		5.90 (3.9)	6.12 (3.8)	7.01 (3.6)	7.48 (4.2)	8.92 (4.5)	<0.001
HADS Depression (mean (SD)) (n= 5,528)		3.49 (3.0)	3.88 (3.0)	4.63 (3.3)	5.25 (3.8)	6.74 (3.7)	<0.001
ACR Widespread pain (n= 5,617)		706 (16.3)	105 (46.9)	141 (45.9)	227 (57.6)	246 (69.5)	<0.001
BMI (mean (SD)) (n= 5,468)		26.6 (4.1)	26.9 (4.0)	27.1 (5.3)	27.9 (6.2)	28.0 (5.1)	<0.001
SF-12 general health (n= 5,562)	Excellent	287 (6.7)	9 (4.0)	4 (1.3)	8 (2.1)	3 (0.9)	<0.001
	Very good	1428 (33.2)	43 (19.3)	67 (22.1)	53 (13.7)	22 (6.3)	
	Good	1853 (43.1)	116 (52.0)	139 (45.9)	168 (43.4)	103 (29.5)	
	Fair	659 (15.3)	53 (23.8)	84 (27.7)	133 (34.4)	153 (43.8)	
	Poor	73 (1.7)	2 (0.9)	9 (3.0)	25 (6.5)	68 (19.5)	
Freq-uency of GP visit (n= 5,593)	Very often	41 (1.0)	1 (0.5)	5 (1.6)	7 (1.8)	15 (4.3)	<0.001
	Often	475 (11.0)	35 (15.6)	67 (21.9)	82 (21.1)	90 (25.6)	
	Occasionally	2251 (52.1)	126 (56.3)	166 (54.3)	239 (61.4)	218 (61.9)	
	Seldom	949 (22.0)	33 (14.7)	50 (16.3)	41 (10.5)	21 (6.0)	
	Hardly ever	606 (14.0)	29 (13.0)	18 (5.9)	20 (5.1)	8 (2.3)	
Sleep a (n= 5,522)	No	1999 (46.9)	86 (38.7)	98 (32.7)	109 (28.1)	78 (22.4)	<0.001
	Some nights	1896 (44.5)	113 (50.9)	160 (53.3)	212 (54.6)	162 (46.1)	
	Most nights	368 (8.6)	23 (10.4)	42 (14.0)	67 (17.3)	109 (31.2)	
Sleep b (n= 5,515)	No	871 (20.5)	34 (15.3)	35 (11.2)	29 (7.5)	19 (5.4)	<0.001
	Some nights	2356 (55.4)	121 (54.5)	144 (47.5)	200 (51.8)	146 (41.6)	
	Most nights	1023 (24.1)	67 (30.2)	125 (41.3)	157 (40.7)	186 (53.0)	

Baseline factor		LA	HP	PGF	HPPGF	SA	*p*-value
Sleep c (n= 5,445)	No	1618 (38.5)	68 (30.8)	77 (26.0)	65 (17.1)	51 (15.0)	<0.001
	Some nights	1987 (47.2)	113 (51.1)	142 (48.0)	208 (54.6)	150 (44.3)	
	Most nights	603 (14.3)	40 (18.1)	77 (26.0)	108 (28.4)	138 (40.7)	
Sleep d (n= 5,509)	No	1903 (44.8)	86 (38.6)	80 (26.4)	88 (22.7)	58 (16.8)	<0.001
	Some nights	1896 (44.6)	104 (46.6)	170 (56.1)	196 (50.7)	160 (46.2)	
	Most nights	451 (10.6)	33 (14.8)	53 (17.5)	103 (26.6)	128 (37.0)	

Footnote: HADS= Hospital Anxiety and Depression Scale, range 0-21, where 0 is no presence of anxiety/ depression, 21 is strong presence of anxiety/ depression (Zigmond and Snaith, 1983); ACR= American College of Rheumatology (Wolfe et al., 1990); SF-12= Short Form questionnaire (Ware et al., 1996);
Sleep a: 'Trouble falling asleep', Sleep b: 'Wake up several times in night', Sleep c: 'Trouble staying asleep', Sleep d: 'Wake up and feel tired'; identification of one of four sleep items 'on most nights' indicate sleep problems (Jenkins et al., 1988).

6.3 LTA model

6.3.1 Interpretation of base model: transitions at 3 and 6 years follow-up

In *Chapter 5*, the item-response probabilities in the base model were restricted to ensure a more straightforward interpretation of the latent states at each time point. Further interpretation of the base model is addressed here, by the assessment of the full LTA model (restricted base model), including the transitional probabilities (i.e. the probability of moving from one state to another over time). The estimated proportions in each latent state, the item-response probabilities and the transition probabilities in the restricted model are shown in *Table 6.3.1*. The estimated proportions in each latent state were very similar to the actual proportion allocated into each state.

The initial three rows of probabilities (latent state proportions) in the table represent the proportion of individuals in each state at each time point. The next eight rows of probabilities are the item-response probabilities (with 95% Confidence Intervals). The latent state proportions at each time point, and the item-response probabilities for the restricted model were very similar to the unrestricted model shown in *Table 4.8.1*.

For this restricted model, there were two sets of transitional probabilities (τ); the probabilities of transitioning between baseline and 3 years, and the probabilities of transitioning between 3 years and 6 years. In *Table 6.3.1*, the rows for latent transition probabilities represent the earlier time point, and the column the next time point. For example, in the top right corner, 0.018 (in *italics* in *Table 6.3.1*) represented the probability of transitioning to 'severely affected' at 3 years follow-up for those in the 'least affected' state at baseline.

Of those in the 'least affected' state at baseline, an estimated 87% remained in the 'least affected' state at 3 years (τ= 0.87). There was little progression to the other states. This pattern was also true from 3 to 6 years. For individuals in the 'high pain' state at baseline, less than half (τ= 0.38) were estimated to remain in this state at 3 years, and a similar probability were estimated to transition back to the 'least affected' state. About one in eight individuals were estimated to progress to the 'high pain and poor gross function' state (τ= 0.15). Between 3 and 6 years, only a quarter stayed in the 'high pain' state (τ= 0.26), with just under half estimated to transition to the 'least affected' state (τ= 0.48).

Of those in the 'poor gross function' state at baseline, just over a quarter were estimated to remain in this state at 3 years (τ= 0.27), with a third of people also estimated to develop high pain issues by 3 years and therefore estimated to transition to the 'high pain and poor gross function' state (τ= 0.33). About a quarter were estimated to improve their hand symptoms and transition into the 'least affected' state (τ= 0.24). The probabilities were similar between 3 and 6 years, however with an increased rate of stability in this state (from 0.27 to 0.35).

Table 6.3.1: LTA parameters for the final base model.

	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Latent state proportions					
Baseline (Time 1)	0.768	0.043	0.058	0.068	0.063
3 years (Time 2)	0.721	0.059	0.047	0.095	0.079
6 years (Time 3)	0.702	0.057	0.046	0.094	0.101
Item-response probabilities (with 95% CI)					
Pain when turning objects	0.001 (0.00, 0.00)	0.733 (0.68, 0.79)	0.125 (0.09, 0.16)	0.915 (0.89, 0.94)	0.977 (0.97, 0.99)
Pain when squeezing objects	0.004 (0.00, 0.01)	0.818 (0.77, 0.87)	0.156 (0.11, 0.20)	0.960 (0.94, 0.98)	0.989 (0.98, 1.00)
Pain when gripping objects	0.006 (0.01, 0.01)	0.763 (0.72, 0.81)	0.146 (0.11, 0.18)	0.865 (0.84, 0.89)	0.973 (0.96, 0.99)
Difficulty opening a new jar	0.005 (0.00, 0.01)	0.228 (0.18, 0.28)	0.728 (0.67, 0.79)	0.897 (0.87, 0.93)	1.000 (1.00, 1.00)
Difficulty carrying a full pot	0.005 (0.00, 0.01)	0.091 (0.05, 0.14)	0.631 (0.57, 0.70)	0.820 (0.78, 0.86)	0.993 (0.99, 1.00)
Difficulty wringing out a dishcloth	0.002 (0.00, 0.00)	0.180 (0.14, 0.22)	0.445 (0.39, 0.50)	0.787 (0.75, 0.83)	0.988 (0.98, 1.00)
Difficulty doing-up buttons	0.001 (0.00, 0.00)	0.038 (0.02, 0.06)	0.172 (0.13, 0.21)	0.238 (0.19, 0.29)	0.917 (0.89, 0.95)
Difficulty turning taps on	0.000 (0.00, 0.00)	0.013 (0.00, 0.02)	0.093 (0.07, 0.12)	0.161 (0.12, 0.20)	0.889 (0.85, 0.93)
Latent transition probabilities					
Baseline to 3 years	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Least affected	**0.867**	0.049	0.026	0.040	*0.018*
High pain	0.417	**0.384**	0.027	0.151	0.021
Poor gross function	0.244	0.031	**0.274**	0.329	0.122
High pain & poor gross function	0.207	0.037	0.094	**0.452**	0.211
Severely affected	0.134	0.006	0.059	0.117	**0.684**
3 years to 6 years	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Least affected	**0.868**	0.048	0.023	0.038	0.023
High pain	0.481	**0.262**	0.031	0.173	0.053
Poor gross function	0.284	0.000	**0.351**	0.273	0.091
High pain & poor gross function	0.222	0.076	0.089	**0.416**	0.198
Severely affected	0.177	0.000	0.034	0.057	**0.733**

Footnote: CI= Confidence Intervals; Bold entries represent stability between time points.

Of individuals classified in the 'high pain and poor gross function' at baseline, just under half were estimated to remain in this state at 3 years (τ= 0.45), one in five to transition into the 'least affected' state (τ= 0.21), and a similar probability of individuals to the 'severely affected' state (τ= 0.21); there was little estimated movement into the 'high pain' or 'poor

gross function' states (τ= 0.04 and τ= 0.09 respectively). All transitions were similar between baseline and 3 years, and 3 years and 6 years. Over two thirds (τ= 0.68) of individuals in the 'severely affected' state at baseline were estimated to remain in this state at 3 years, increasing to nearly three quarters (τ= 0.73) between 3 years and 6 years. Of those 'severely affected' at baseline, roughly one in nine individuals were estimated to experience a slight improvement in symptoms and transition into the 'high pain and poor gross function' state (τ= 0.12), and about one in eight individuals into 'least affected' at 3 years (τ= 0.13). This transition between 'severely affected' and 'least affected' was increased between 3 years and 6 years, from 0.13 to 0.18 (nearly one in five).

6.3.2 6 year transitions

While *Table 6.3.1* displayed the transitional patterns of baseline to 3 years, then 3 years to 6 years, this presentation did not reveal the full baseline to 3 years to 6 years pattern. This may mask the fact that some individuals returned to their baseline state at 6 years (after transitioning out at 3 years). *Table 6.3.2* displayed this pattern and accounted for all potential pathways (five latent states at three time points= 5^3= 125). An alternative presentation in the format of flow diagrams is presented in *Appendix D*. Every cell in *Table 6.3.2* represented the number in a 6 year state stratified by baseline and 3 year state.

For example, of the 3,777 individuals who were in 'least affected' at both baseline and at 3 years, 3,371 remained in 'least affected' at 6 years (78% of 'least affected' at baseline). An example of individuals returning to their baseline state could be seen in the first section of the table, where individuals in 'least affected' at baseline, then transitioned into 'high pain' at 3 years, 108 (2% of those 'least affected' at baseline, or 58% of the 195 in 'high pain' at 3 years and 'least affected' at baseline) then returned back to 'least affected' at 6 years; more than double the number of people that remained in 'high pain' at 6 years.

Another pattern to highlight was that there did appear to be little transitioning into the 'high pain' group at 6 years from any group that was not 'high pain' at 3 years (and a similar pattern was evident from the transition probabilities in *Table 6.3.1*). Similar to this, there was little transition into the 'poor gross function' state at 6 years. A further exploration of the baseline to 6 year pattern is displayed in *Appendix E*, where the LTA approach is applied excluding the 3 years follow-up point. Overall, the transition probabilities of the main analysis of baseline to 3 years, and 3 years to 6 years (*Table 6.3.1*) were similar to those for baseline to 6 years transition probabilities.

Table 6.3.2: Transitions between baseline, 3 years and 6 years, for all individuals.

Baseline	3 years	6 years				
		LA	HP	PGF	HPPGF	SA
Least affected (n= 4,338)	LA (n= 3,777)	3,371 (78%)	171 (4%)	63 (1%)	130 (3%)	42 (1%)
	HP (n= 195)	108 (2%)	45 (1%)	3 (<1%)	30 (1%)	9 (<1%)
	PGF (n= 105)	34 (1%)	0 (0%)	34 (1%)	34 (1%)	3 (<1%)
	HPPGF (n= 181)	60 (1%)	17 (<1%)	14 (<1%)	61 (1%)	29 (1%)
	SA (n= 80)	27 (1%)	0 (0%)	4 (<1%)	5 (<1%)	44 (1%)
High pain (n= 224)	LA (n= 90)	63 (28%)	14 (6%)	1 (<1%)	9 (4%)	3 (1%)
	HP (n= 100)	36 (16%)	35 (16%)	2 (1%)	23 (10%)	4 (2%)
	PGF (n= 4)	0 (0%)	0 (0%)	2 (1%)	2 (1%)	0 (0%)
	HPPGF (n= 27)	7 (3%)	2 (1%)	1 (<1%)	14 (6%)	3 (1%)
	SA (n= 3)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (1%)
Poor gross function (n= 307)	LA (n= 65)	44 (14%)	3 (1%)	5 (2%)	8 (3%)	5 (2%)
	HP (n= 7)	4 (1%)	1 (<1%)	0 (0%)	2 (1%)	0 (0%)
	PGF (n= 84)	22 (7%)	0 (0%)	38 (12%)	18 (6%)	6 (2%)
	HPPGF (n= 110)	17 (6%)	2 (1%)	14 (5%)	65 (21%)	12 (4%)
	SA (n= 41)	5 (2%)	0 (0%)	5 (2%)	1 (<1%)	30 (10%)
High pain & poor gross function (n= 394)	LA (n= 95)	51 (13%)	5 (1%)	5 (1%)	21 (5%)	13 (3%)
	HP (n= 9)	4 (1%)	2 (1%)	0 (0%)	2 (1%)	1 (0%)
	PGF (n= 36)	9 (2%)	0 (0%)	9 (2%)	14 (4%)	4 (1%)
	HPPGF (n= 182)	33 (8%)	12 (3%)	17 (4%)	98 (25%)	22 (6%)
	SA (n= 72)	8 (2%)	0 (0%)	3 (1%)	4 (1%)	57 (14%)
Severely affected (n= 354)	LA (n= 41)	16 (5%)	0 (0%)	4 (1%)	9 (3%)	12 (3%)
	HP (n= 2)	0 (0%)	1 (<1%)	0 (0%)	1 (<1%)	0 (0%)
	PGF (n= 17)	3 (1%)	0 (0%)	2 (1%)	8 (2%)	4 (1%)
	HPPGF (n= 35)	5 (1%)	2 (1%)	0 (0%)	13 (4%)	15 (4%)
	SA (n= 259)	39 (11%)	0 (0%)	5 (1%)	14 (4%)	201 (57%)

Footnote: LA= 'least affected'; HP= 'high pain'; PGF= 'poor gross function'; HPPGF= 'high pain and poor gross function'; SA= 'severely affected'; %'s are of state at baseline.

6.4 Stratified transition probabilities

In the previous sections, the LTA process had followed the standard approach for classifying individuals into latent states and then investigating how individuals transition between states over the three time points, but as one whole population. There were two additional approaches which can be used to explore whether transitional probabilities were affected by respondent characteristics. The first approach permits the transition probabilities to differ by various characteristics (the approach taken in this section, *section 6.4*), whereas the second uses the population as a whole but investigates the influence of a covariate on the estimated transition probabilities (*section 6.5*). It is worth noting that in addition to this, factors can be used to predict future state membership, which is discussed and used throughout *Chapter 7*.

In this first exploratory approach, four demographic and pain characteristics were assessed, which through discussions with supervisors, and taking on board comments made in the RUG meeting and from the literature, were thought most likely to influence transition probabilities.

6.4.1 Methods

The LTA model was computed again based on five latent states with the same eight indicators and restricted item-response probabilities over time, but this time generating separate transition probabilities for different characteristics. For example, in the first case of gender, males and females each had their own transitional matrices for baseline (time 1) to 3 years (time 2), and 3 years (time 2) to 6 years (time 3). In this model, the transitional probabilities (τ's) were permitted to vary by gender, whilst the other parameters were estimated for the population as a whole (such as state proportions and item-response probabilities).

This “unrestrained” model was then compared to the previously derived restricted base model displayed in *Table 6.3.1*, which contained “restraints” on the transitional probabilities (τ's), i.e. forcing males and females to have the same transitional pattern. Therefore, the only difference between these two models was that the “unrestrained” model allowed the transition probabilities to be estimated separately for males and females; all other parameters were estimated using the population as a whole.

A LRT between the two models was used to compare model fit between the “restrained” and “unrestrained” models. In addition to this, the BIC values of each model were compared, with the model with the lower BIC preferred. Furthermore, the transitional probabilities in the “unrestrained” model were visually inspected between males and females to see if they appeared comparable.

Four characteristics were assessed in separate “unrestrained” models. Firstly the possibility that transition probabilities were different between males and females was explored. Gender is one of the more common characteristics related to the development of hand pain/ problems in older people with females more likely to develop more severe hand OA (van Saase et al., 1989; Zhang et al., 2002; Aihie Sayer et al., 2003; Zhang et al., 2003; Bellamy et al., 2011; Kwok, 2013; Visser et al., 2014). Based on this, and discussions with the supervisors and RUG, gender was highlighted as one of the main characteristics worth exploring to assess whether males have a different transitional pattern to females.

Secondly, whether individuals lived alone or not was considered. As explained in *section 4.4.3*, the RUG meeting allowed the opportunity for individuals with hand pain/ problems/ OA to give their own history and experiences of living with the condition. One aspect that came out of the discussion was patients who lived alone reported finding it harder to cope with the condition. The potential reasons were not having others around to help with the

daily activities that hand pain/ problems affected, such as assistance with 'difficulty opening a new jar' or 'difficulty doing-up buttons'. Therefore, living status was investigated in this section to explore whether baseline living status affected the transitional pattern of hand problems.

Thirdly, age is widely considered as the most influential characteristic of the development and progression of hand pain/ problems (Caspi et al., 2001; Haara et al., 2003; Grotle et al., 2008a; Ghosh et al., 2014; Prieto-Alhambra et al., 2014). Therefore, a third model assessed whether the transitional patterns varied between different age groups. The age groups of focus were 50-64 years, to represent the younger 'pre-pension' age, 65-74 years to represent a 'newly retired' age, and 75+ years to represent the 'oldest' age group.

Finally, the transitional patterns of individuals with and without widespread pain were compared. Individuals were classified as having widespread pain based on the pain manikin included in the HS questionnaire, using the ACR criteria (Wolfe et al., 1990). It has been suggested by previous literature that individuals with widespread pain are more likely to be affected by hand OA (Grotle et al., 2008a; Nicholls et al., 2012), and therefore it was of interest to explore whether individuals were associated with the progression of hand pain/ problems in this study. This analysis was computed using Mplus version 7.3 (Muthén and Muthén, 1998-2015).

6.4.2 Results - Gender

Of the study population, 54% (n= 3,031) were female and 46% (n= 2,586) were male (as shown previously in *Table 4.7.1*). *Table 6.4.1* showed that 9% of females (n= 261), but only 4% of males (n= 93) were in the 'severely affected' state at baseline in the base model. There was a similar pattern in the 'high pain and poor gross function' and the 'poor

gross function' states. However, a higher proportion of the males than of the females were in the 'least affected' and 'high pain' states.

In the unrestrained model, which allowed transition probabilities to vary by gender, there was a noticeable difference between males and females in the estimated stability within the 'poor gross function' state. Between baseline to 3 years, males had a lower probability (τ= 0.102) of remaining in this state, compared to females (τ= 0.297). However, between 3 and 6 years, the probability for males remaining in the 'poor gross function' state increased to 0.360, consistent with the 0.342 probability of the females. Similarly, males were less likely to remain in the 'severely affected' group (τ= 0.550 for males vs. τ= 0.737 for females, baseline to 3 years), and hence had a higher probability of improvement in their hand symptoms 3 years later.

Females had a lower chance of remaining in the 'least affected' state in each transition period indicating a higher probability of deterioration 3 years later. Although the differences were not large (0.833 vs. 0.900 for baseline to 3 years, 0.843 vs. 0.892 for 3 years to 6 years), considering the frequency of 'least affected' at each time point, this did represent a distinct difference for a large number of people.

Table 6.4.1: Latent state proportions and transition probabilities for females and males (females always top entry).

Gender	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Baseline latent state proportions					
Female (54.0%)	0.704	0.031	0.092	0.088	0.085
(n= 3,031)	(n= 2,167)	(n= 96)	(n= 241)	(n= 266)	(n= 261)
Male (46.0%)	0.839	0.062	0.022	0.042	0.036
(n= 2,586)	(n= 2,171)	(n= 128)	(n= 66)	(n= 128)	(n= 93)
Latent transition probabilities					
Baseline to 3 years	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Least affected					
(Female)	**0.833**	0.042	0.042	0.057	0.026
(Male)	**0.900**	0.057	0.012	0.020	0.011
High pain					
(Female)	0.426	**0.371**	0.022	0.159	0.022
(Male)	0.401	**0.385**	0.054	0.125	0.035
Poor gross function					
(Female)	0.270	0.026	**0.297**	0.304	0.103
(Male)	0.232	0.102	**0.102**	0.406	0.158
High pain & poor gross function					
(Female)	0.212	0.012	0.106	**0.456**	0.214
(Male)	0.176	0.105	0.049	**0.461**	0.209
Severely affected					
(Female)	0.114	0.003	0.052	0.094	**0.737**
(Male)	0.179	0.034	0.086	0.152	**0.550**
3 years to 6 years	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Least affected					
(Female)	**0.843**	0.041	0.028	0.055	0.033
(Male)	**0.892**	0.056	0.020	0.019	0.013
High pain					
(Female)	0.439	**0.236**	0.061	0.240	0.024
(Male)	0.486	**0.277**	0.023	0.120	0.094
Poor gross function					
(Female)	0.328	0.000	**0.342**	0.261	0.069
(Male)	0.145	0.047	**0.360**	0.298	0.150
High pain & poor gross function					
(Female)	0.210	0.047	0.101	**0.466**	0.177
(Male)	0.240	0.163	0.065	**0.298**	0.234
Severely affected					
(Female)	0.151	0.005	0.029	0.034	**0.780**
(Male)	0.253	0.000	0.041	0.112	**0.595**

Footnote: n= number of observations; Bold entries represent stability between time points between time points.

Table 6.4.2 summarises the comparison between the two models, i.e. the base model restraining the transitional probabilities to be the same between males and females and the new model assuming they were different ("unrestrained" model). Firstly, the BIC was lower for the base model, implying that this model fitted better than the unrestrained model. However, when comparing the likelihoods through the LRT, the unrestrained model was a significant improvement on the base model, therefore implying that the transitional patterns were significantly different between males and females. In conclusion, there were visual disparities between the transitional probabilities of males and females, with females more likely to remain in the 'severely affected' state at 3 and 6 years.

Table 6.4.2: LTA model comparisons for differences between genders.

	Tau restrained model	Tau unrestrained model
BIC	58,892.339	59,035.540
Log-likelihood	-29,061.977	-28,960.906
Degrees of freedom	89	129
Difference in log-likelihood (×2)		202.142
Difference in degrees of freedom (×2)		80
p-value		<0.001

Footnote: BIC= Bayesian Information Criteria.

6.4.3 Results - Living status

Approximately 19% of the study population (n= 1,020) reported living alone at baseline, and a higher proportion of those living alone (9%, n= 94) were classified in the most severe state at baseline compared to those who stated they did not live alone (5%, n= 238) in the base model (*Table 6.4.3*).

Table 6.4.3: Latent state proportions and transition probabilities for individuals that did and did not live alone (lived alone always top entry).

Living status	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Baseline latent state proportions					
Lived alone (18.9%) (n= 1,020)	0.729 (n= 752)	0.033 (n= 29)	0.067 (n= 60)	0.080 (n= 85)	0.091 (n= 94)
Did not live alone (81.1%) (n= 4,388)	0.777 (n= 3,426)	0.046 (n= 185)	0.057 (n= 239)	0.067 (n= 300)	0.053 (n= 238)
Latent transition probabilities					
Baseline to 3 years	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Least affected					
(live alone)	**0.842**	0.044	0.29	0.048	0.036
(do not live alone)	**0.874**	0.052	0.026	0.037	0.011
High pain					
(live alone)	0.404	**0.416**	0.034	0.117	0.030
(do not live alone)	0.402	**0.379**	0.025	0.168	0.027
Poor gross function					
(live alone)	0.334	0.000	**0.239**	0.343	0.084
(do not live alone)	0.215	0.048	**0.286**	0.319	0.132
High pain & poor gross function					
(live alone)	0.145	0.000	0.118	**0.446**	0.290
(do not live alone)	0.226	0.055	0.084	**0.457**	0.178
Severely affected					
(live alone)	0.198	0.000	0.044	0.091	**0.667**
(do not live alone)	0.105	0.007	0.058	0.141	**0.689**
3 years to 6 years	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Least affected					
(live alone)	**0.874**	0.038	0.029	0.035	0.025
(do not live alone)	**0.866**	0.051	0.022	0.039	0.022
High pain					
(live alone)	0.524	**0.249**	0.000	0.207	0.020
(do not live alone)	0.474	**0.254**	0.037	0.178	0.057
Poor gross function					
(live alone)	0.275	0.041	**0.290**	0.324	0.072
(do not live alone)	0.297	0.000	**0.369**	0.255	0.079
High pain & poor gross function					
(live alone)	0.243	0.021	0.063	**0.497**	0.176
(do not live alone)	0.216	0.095	0.094	**0.393**	0.201
Severely affected					
(live alone)	0.186	0.000	0.000	0.053	**0.761**
(do not live alone)	0.165	0.000	0.046	0.062	**0.727**

Footnote: n= number of observations; Bold entries represent stability between time points.

In general, there were few differences in the transitional probabilities according to living status. In the baseline to 3 years transitions, the main differences were that a higher proportion of individuals who did not live alone transitioned from 'high pain and poor gross function' into 'least affected' (0.226 vs. 0.145), and a higher proportion of people who did live alone transitioned into 'severely affected' from 'high pain and poor gross function' (0.290 vs. 0.178). These results implied less recovery and more deterioration in those who lived alone compared to those who did not. There were differences between the stability of states between 3 and 6 years (namely 'poor gross function' and 'high pain and poor gross function'), however these differences were not so evident in the baseline to 3 year period. This effect could potentially be due to differences in age between those who lived alone and those who did not. The effect of age is explored in the next section.

In addition to the visual inspection of the transitional probabilities for living status, *Table 6.4.4* showed that the BIC was lower in the base model with the restrained tau (τ) probabilities (same transition probabilities for each living status). In addition to this, the LRT was also non-significant (p= 0.977). Therefore, there appeared to be no clear evidence of a noticeable or significant difference in the transitional patterns between individuals with hand pain/ problems according to living status.

Table 6.4.4: LTA model comparisons for differences between living status.

	Tau restrained model	Tau unrestrained model
BIC	54,781.093	55,068.064
Log-likelihood	-27,008.041	-26,979.613
Degrees of freedom	89	129
Difference in log-likelihood (×2)		56.856
Difference in degrees of freedom (×2)		80
p-value		0.977

Footnote: BIC= Bayesian Information Criteria.

6.4.4 Results - Age

At baseline, the older the individual, the more likely they were of being classified in the more severe hand states (as shown previously in *Table 6.2.1*). Comparing across age brackets, there were many differences in transitional probabilities when allowing the probabilities to vary by age group (*Table 6.4.5*). One of the key findings was that the stability probabilities for the older age group were lower for nearly every state compared to the younger age groups. In addition to this, while all age groups demonstrated an increased probability for transitioning into the 'least affected' state from 3 years to 6 years, these probabilities were higher for the oldest age group (75+ years) than the younger age groups for those in states containing high function problem indicators (for example, 'severely affected' at 3 years to 'least affected' at 6 years, τ= 0.147 for 50-64 years, τ= 0.186 for 65-74 years, τ= 0.225 for 75+ years). Therefore, the older age groups apparently showed more recovery into the 'least affected' state from high pain and poor function states compared to the younger groups.

This finding was converse to what was expected based on previous literature. However, it was important to note that the oldest age group had the smallest number of individuals (about 10% (n= 540) of the population), with only 44 of these in the 'severely affected' state at baseline, so these probabilities should be interpreted with caution as they reflected a smaller number of observations.

Table 6.4.5: Latent state proportions and transition probabilities for individuals aged 50 years and over at baseline (50-64 always top entry, 75+ always bottom entry).

Age group (years)	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Latent state proportions					
50-64 (60.9%)	0.787	0.047	0.048	0.066	0.051
(n= 3,418)	(n= 2,700)	(n= 147)	(n= 160)	(n= 235)	(n= 176)
65-74 (29.5%)	0.734	0.041	0.071	0.075	0.080
(n= 1,659)	(n= 1,231)	(n= 60)	(n= 110)	(n= 124)	(n= 134)
75+ (9.6%)	0.746	0.036	0.079	0.053	0.085
(n= 540)	(n= 407)	(n= 17)	(n= 37)	(n= 35)	(n= 44)
Latent transition probabilities					
Baseline to 3 years	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Least affected					
(50-64)	**0.873**	0.048	0.026	0.039	0.014
(65-74)	**0.865**	0.047	0.019	0.044	0.025
(75+)	**0.834**	0.055	0.038	0.047	0.025
High pain					
(50-64)	0.377	**0.439**	0.017	0.147	0.021
(65-74)	0.455	**0.313**	0.085	0.104	0.043
(75+)	0.566	**0.138**	0.000	0.296	0.000
PGF					
(50-64)	0.222	0.031	**0.295**	0.343	0.109
(65-74)	0.234	0.044	**0.280**	0.367	0.075
(75+)	0.360	0.000	**0.214**	0.080	0.346
High pain & PGF					
(50-64)	0.171	0.028	0.091	**0.537**	0.172
(65-74)	0.250	0.080	0.054	**0.365**	0.252
(75+)	0.292	0.000	0.196	**0.156**	0.355
Severely affected					
(50-64)	0.095	0.012	0.076	0.127	**0.691**
(65-74)	0.180	0.000	0.032	0.086	**0.703**
(75+)	0.146	0.000	0.101	0.169	**0.584**
3 years to 6 years	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Least affected					
(50-64)	**0.872**	0.057	0.018	0.033	0.020
(65-74)	**0.861**	0.040	0.024	0.045	0.031
(75+)	**0.864**	0.021	0.046	0.048	0.021
High pain					
(50-64)	0.478	**0.266**	0.042	0.173	0.042
(65-74)	0.487	**0.269**	0.023	0.124	0.097
(75+)	0.412	**0.181**	0.000	0.347	0.060
PGF					
(50-64)	0.284	0.020	**0.339**	0.248	0.109
(65-74)	0.270	0.000	**0.378**	0.351	0.000
(75+)	0.317	0.000	**0.336**	0.258	0.090
High pain & PGF					
(50-64)	0.204	0.093	0.086	**0.449**	0.168
(65-74)	0.203	0.032	0.093	**0.434**	0.238
(75+)	0.513	0.079	0.091	**0.121**	0.196
Severely affected					
(50-64)	0.147	0.007	0.032	0.062	**0.752**
(65-74)	0.186	0.000	0.038	0.067	**0.709**
(75+)	0.225	0.000	0.015	0.000	**0.760**

Footnote: n= number of observations; Bold entries represent stability between time points; PGF= Poor gross function.

Investigating the model summary data in *Table 6.4.6*, the comparison of BIC's between the two models suggested that the unrestrained model allowing transition probabilities to vary by age group was more favourable. However, the *p*-value from the LRT produced a non-significant outcome, demonstrating that there was no significant difference between the models, and therefore between the transitional patterns across age groups. Therefore, this investigation had contradictory results between the BIC and the LRT. It was evident that there were some differences in transitional patterns, however some probabilities could have been influenced by small state sizes.

Table 6.4.6: LTA model comparisons for differences between ages.

	Tau restrained model	Tau unrestrained model
BIC	61,369.864	60,773.903
Log-likelihood	-30,279.155	-30,212.951
Degrees of freedom	94	174
Difference in log-likelihood (×2)		132.408
Difference in degrees of freedom (×2)		160
p-value		0.946

Footnote: BIC= Bayesian Information Criteria.

6.4.5 Results - Widespread pain

Just over a quarter of the population were classified as having widespread pain at baseline (25.4%). In the base model individuals with widespread pain were at least twice as likely to belong to any of the problematic hand states, and almost seven times more likely to be classified in the 'severely affected' state at baseline (17% vs. 3%) (*Table 6.4.7*).

Table 6.4.7: Latent state proportions and transition probabilities for individuals with (WP) and without widespread pain (No WP) at baseline.

Widespread pain	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Baseline latent state proportions					
Widespread pain (WP) (25.4%) (n= 1,425)	0.480 (n= 706)	0.082 (n= 105)	0.109 (n= 141)	0.156 (n= 227)	0.174 (n= 246)
No widespread pain (No WP) (74.6%) (n= 4,192)	0.865 (n= 3,632)	0.030 (n= 119)	0.043 (n= 166)	0.037 (n= 167)	0.025 (n= 108)
Latent transition probabilities					
Baseline to 3 years	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Least affected					
(WP)	**0.750**	0.087	0.048	0.074	0.041
(No WP)	**0.888**	0.041	0.023	0.033	0.015
High pain					
(WP)	0.327	**0.454**	0.010	0.209	0.000
(No WP)	0.505	**0.316**	0.048	0.098	0.032
Poor gross function					
(WP)	0.229	0.000	**0.223**	0.403	0.145
(No WP)	0.281	0.060	**0.311**	0.254	0.094
High pain & poor gross function					
(WP)	0.210	0.024	0.100	**0.484**	0.183
(No WP)	0.206	0.041	0.088	**0.402**	0.264
Severely affected					
(WP)	0.072	0.011	0.068	0.105	**0.745**
(No WP)	0.287	0.000	0.043	0.148	**0.523**
3 years to 6 years	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Least affected					
(WP)	**0.757**	0.073	0.050	0.065	0.055
(No WP)	**0.889**	0.043	0.018	0.033	0.017
High pain					
(WP)	0.417	**0.305**	0.000	0.213	0.065
(No WP)	0.527	**0.226**	0.047	0.154	0.046
Poor gross function					
(WP)	0.158	0.045	**0.312**	0.323	0.163
(No WP)	0.370	0.000	**0.370**	0.219	0.042
High pain & poor gross function					
(WP)	0.223	0.049	0.106	**0.393**	0.228
(No WP)	0.220	0.095	0.070	**0.445**	0.170
Severely affected					
(WP)	0.150	0.000	0.028	0.063	**0.759**
(No WP)	0.222	0.018	0.042	0.035	**0.684**

Footnote: n= number of observations; WP= Widespread pain (based on ACR (American College of Rheumatology) widespread pain (Wolfe et al., 1990); Bold entries represent stability between time points.

There were many differences in transition probabilities between participants with and without widespread pain. In general, transitioning to (or remaining in) the 'least affected' state at either 3 or 6 years was much more likely in the individuals without widespread pain (all but one of the transition probabilities to 'least affected' state were greater in the no widespread pain group). Conversely, there was a much higher probability of transitioning to (or remaining in) the 'high pain and poor gross function' state if the individual had widespread pain.

Another finding to highlight was that those who had widespread pain in the 'poor gross function' state at either baseline or 3 years were more likely to deteriorate to the 'high pain and poor gross function' or 'severely affected' at the next time point (baseline to 3 years/ 3 years to 6 years: widespread pain: 0.403/ 0.323 to 'high pain and poor gross function', 0.145/ 0.163 to 'severely affected', and no widespread pain: 0.254/ 0.219 to 'high pain and poor gross function', 0.094/ 0.042 to 'severely affected'). This finding implied that individuals with widespread pain were more likely to move towards a more painful hand condition if they already had functional issues, and therefore those with widespread pain were more likely to have poorer outcomes.

Comparison between the "restrained" base and "unrestrained" models for widespread pain revealed that the BIC was lower in the base model (*Table 6.4.8*). The LRT indicated however, that there was a significant difference in the transitional patterns between those with and without widespread pain. Considering the information gained previously from the differences in widespread pain between the latent states (*Table 6.2.1*), there was some evidence that widespread pain did vary between states, and transitions with more deterioration and less recovery were observed in people with widespread pain. This hypothesis is tested further in the next chapter where widespread pain as a predictive factor of future state membership is assessed.

Table 6.4.8: LTA model comparisons for differences between individuals with and without widespread pain.

	Tau restrained model	Tau unrestrained model
BIC	56,916.073	57,019.981
Log-likelihood	-28,073.843	-27,953.126
Degrees of freedom	89	129
Difference in log-likelihood (×2)	241.43	
Difference in degrees of freedom (×2)	80	
p-value	<0.001	

Footnote: BIC= Bayesian Information Criteria.

6.5 Inclusion of covariates in latent state development

6.5.1 Methods

An alternative approach to assess how covariates influence state membership and state transitions is to include these factors directly into the modelling process when the latent states are computed. In *section 6.4*, the model was assessed using the whole population but an investigation of whether transitions varied by a characteristic of interest (stratified analysis) was then performed. In this section, these characteristics were included in the model as an adjustment (termed covariates for easier reading), so individuals were classified into their latent states based on the indicators and also the adjusted covariate. There are some parallels between this analysis and the transitional split analysis in *section 6.4* in that both explored differences between, for example, males and females. However, the analysis presented in *section 6.4* aimed to visually assess (and test for) differences in transitional probabilities between males and females, whereas the subsequent analysis explored whether it was necessary to adjust for gender in the latent state development.

The investigation used in this section incorporated confounders as covariates into the modelling process. For example, in the case of gender, gender was incorporated into the state development, not to be used as an indicator, but as an adjustment in state development. The same indicators defined in *section 4.8* were still used to create the states, but the inclusion of gender as a confounder corrected for the possibility that females, for example, may be both more prevalent in a particular state, and prevalent in particular indicators. Gender was still permitted to vary between states, but within states gender was not allowed to be associated with the indicators.

If gender had no effect on the derived model, all the probabilities (state proportions, item-response probabilities and transition probabilities) created during the LTA process should be comparable with the restricted base model (the model displayed in *Table 6.3.1*). However, if gender did have an impact, for example, females were more likely to have function problems and males were more likely to have pain or no problems, the parameters calculated by the LTA process would be different. A more detailed and technical explanation of this process can be found in Muthén and Asparouhov, 2011.

For this investigation, the four factors used previously (gender, age, living status and widespread pain) were included separately as potential confounders. For each of the models, the BIC, entropy and a LRT were reported to indicate if there were any significant differences between the adjusted and the base model. Next, the LTA parameters from the adjusted models (state proportions, item-response probabilities and transition probabilities) were compared against the estimates computed in the base model (*Table 6.3.1*) that was presented in *section 6.3*. In addition to this, the baseline state membership of individuals in each of the adjusted models was compared against the state membership of individuals in the base model. Any clear discrepancies between the estimates in the base model and those in the model with adjusted covariates would indicate a necessity to include the factors in

the modelling process, as this would indicate confounding of LTA estimates by these covariates.

The analysis also produced estimates reflecting the effect of the covariate included on membership in each latent state, whilst adjusting for any previous state membership (for time points after baseline). This analysis was analogous to a standard regression analysis, whilst taking into account any potential error in membership classification. For baseline state, it is the odds of being in a particular latent state depending on the adjusted factor, i.e. does being male predict what latent state you are in at baseline? At future time points, 3 years for example, it is the odds of being in a particular latent state at 3 years, dependent on being male, whilst adjusting for baseline state membership. The odds ratios for the example presented (which is using gender as a covariate) are displayed with a 95% confidence interval.

6.5.2 Results

The BIC, entropy and LRT *p*-values of including each covariate in separate models is presented in *Table 6.5.1*. The LRT was calculated using the formulas displayed in *Chapter 5* (*Equations 5.1* to *5.4*), with each of the covariates increasing the number of degrees of freedom by 12 (five states – 1 ×three time points), and were all in reference to the restricted base model.

There was a decrease in BIC for each of the models (with the exception of age), and a significant difference in model fit using the LRT between the restricted base model, and each of the models with a covariate. However, there were no improvements in entropy for any of the models with a covariate.

Table 6.5.1: Comparison of model results for models with an adjusted factor compared with the standard base model.

Model	BIC	Entropy	Log-likelihood	Degrees of freedom	LRT *p*-value
Restricted base model (from *Table 5.2*)	51,348.727	0.940	-25,311.754	84	n/a
Including as an adjusted factor:					
Gender	51,075.671	0.939	-25,123.425	96	<0.001
Age	51,398.857	0.940	-25,285.018	96	<0.001
Living status	49,589.523	0.940	-24,382.171	96	<0.001
Widespread pain	50,404.784	0.939	-24,787.981	96	<0.001

Footnote: BIC= Bayesian Information Criteria; LRT= Likelihood Ratio Test.

As indicated in the Methods (*section 6.5.1*), it was important to check for estimated parameter differences between the models (model with covariate vs. base model). If there were clear differences, it would indicate a necessity to explore why these differences were present, and potentially to include the covariate in the model for future analysis (or to further employ the stratified analyses as in *section 6.4*). Only the gender estimates are presented here, the estimates from the other covariate models are presented in *Appendix E*.

Comparing the parameter estimates displayed in *Table 6.5.2* (model including gender as a covariate) with those in *Table 6.3.1* (the base model), there were few differences (all differences <0.02). The proportions of people in each latent state and the item-response probabilities were almost identical and there were only a few subtle differences in the latent transition probabilities ('high pain' to 'least affected' at baseline to 3 years, τ= 0.406 in the gender adjusted model and τ= 0.417 in the base model, and, at 3 years to 6 years, τ= 0.465 in the gender adjusted model, and τ= 0.481 in the base model).

Table 6.5.2: LTA estimates when using gender as a predictive factor.

	LA	HP	PGF	HPPGF	SA
Latent state proportions					
Baseline	0.767	0.045	0.059	0.067	0.062
3 years	0.721	0.060	0.048	0.094	0.079
6 years	0.702	0.058	0.047	0.093	0.100
Item-response probabilities					
Pain when turning objects	0.001	0.739	0.125	0.913	0.977
Pain when squeezing objects	0.004	0.825	0.152	0.959	0.988
Pain when gripping objects	0.006	0.777	0.139	0.862	0.973
Difficulty opening a new jar	0.004	0.234	0.724	0.905	1.000
Difficulty carrying a full pot	0.004	0.093	0.623	0.833	0.993
Difficulty wringing out a dishcloth	0.002	0.195	0.434	0.790	0.988
Difficulty doing-up buttons	0.001	0.047	0.159	0.242	0.918
Difficulty turning taps on	0.000	0.016	0.090	0.165	0.892
Latent transition probabilities					
Baseline to 3 years	LA	HP	PGF	HPPGF	SA
Least affected	**0.867**	0.048	0.027	0.040	0.018
High pain	0.406	**0.383**	0.035	0.148	0.028
Poor gross function	0.262	0.036	**0.270**	0.318	0.114
High pain & poor gross function	0.204	0.044	0.090	**0.451**	0.211
Severely affected	0.135	0.009	0.057	0.116	**0.684**
3 years to 6 years	LA	HP	PGF	HPPGF	SA
Least affected	**0.867**	0.048	0.024	0.038	0.023
High pain	0.465	**0.260**	0.037	0.176	0.063
Poor gross function	0.300	0.005	**0.347**	0.266	0.083
High pain & poor gross function	0.221	0.080	0.089	**0.413**	0.196
Severely affected	0.176	0.000	0.034	0.057	**0.734**

Footnote: LA= 'least affected'; HP= 'high pain'; PGF= 'poor gross function'; HPPGF= 'high pain and poor gross function'; SA= 'severely affected'; Bold entries represent stability between time points.

There were also few differences between the base model and covariate model estimates for latent state proportions and item-response probabilities when incorporating the live alone, widespread pain or age covariates (*Appendix E*). In the widespread pain adjusted model, there were some small differences in the 3 year to 6 year transition probabilities. Firstly, in the 'severely affected' state at 3 years, the probability of remaining in the same state at 6 years decreased from 0.733 (in the base model), to 0.686 (in the widespread pain covariate model). In addition to this, again in 'severely affected' at 3 years, the probability of transitioning into 'least affected' at 6 years changed from 0.177 (in the base model), to 0.224 (in the widespread pain covariate model). The results in *section 6.2* demonstrated that individuals in more problematic states were more likely to have widespread pain. Adjusting for widespread pain in state development therefore potentially eliminated some

of the probability of remaining in the 'severely affected' state, and thus permitted an increase in the likelihood of transitioning into the 'least affected' state.

There appeared to be few differences in the classification of individuals between the gender covariate model and the base model. Any small differences were seen in the middle 3 states ('high pain', 'poor gross function' and 'high pain and poor gross function'), which had the lowest numbers of observations. In addition to this, over 98% of individuals that were classified in either 'least affected' or 'severely affected' in the base model, were classified in the same state in the gender adjusted model across all time points. The largest discrepancy was for 20 individuals (3.4% of state) who, at 6 years, were classified as being in 'high pain and poor gross function' in the base model, but were classified as being in the 'high pain' state in the gender adjusted model. None of the other adjusted models had more than 20 individuals that were classified in a different state, the second largest discrepancy was for 10 individuals (1.7% of state proportion) in the widespread pain adjusted model ('high pain and poor gross function' in base model, and 'high pain' in adjusted model) at baseline. All other discrepancies consisted of fewer than seven individuals for all other models. Therefore, the influence of any of these factors in terms of affecting state membership or transitions was small and not sufficient to include adjusted factors in the model.

The odds ratios displayed in *Table 6.5.3* represented the odds of males, compared to females, being classified in each latent state, compared to being classified in the 'least affected' state. For the odds ratios at 3 years (or at 6 years), the estimates have been adjusted for the immediate previous state membership at baseline (or at 3 years). Compared to females, males were more likely to belong to the 'high pain' state instead of the 'least affected' state at baseline (OR= 1.58 (1.16, 2.14)), but significantly less likely to belong to any of the states with function issues (ORs <0.44). This matches the previous

assessments of baseline characteristics in *Table 5.1*. Investigating 3 and 6 year state membership did not alter the observed relationship. The association of males with 'high pain' state was still evident at 3 and 6 years, albeit not significant at the 5% level. In the analysis which included the other covariates (age, living status and widespread pain), the most consistent finding was that individuals that were older, lived alone and had widespread pain respectively, were most likely to be in the 'severely affected' state.

Table 6.5.3: Odds ratios of state membership with gender as a predictor (with adjustment for immediate prior state membership where necessary).

	LA	High pain	Poor gross function	High pain & poor gross function	Severely affected
Baseline					
Female	(ref)	(ref)	(ref)	(ref)	(ref)
Male	(ref)	1.577	0.195	0.442	0.342
		(1.160, 2.143)	(0.134, 0.284)	(0.327, 0.599)	(0.258, 0.454)
3 years					
Female	(ref)	(ref)	(ref)	(ref)	(ref)
Male	(ref)	1.288	0.304	0.563	0.444
		(0.945, 1.755)	(0.198, 0.466)	(0.427, 0.743)	(0.316, 0.624)
6 years					
Female	(ref)	(ref)	(ref)	(ref)	(ref)
Male	(ref)	1.315	0.651	0.482	0.599
		(0.981, 1.763)	(0.454, 0.935)	(0.360, 0.646)	(0.441, 0.814)

Footnote: LA= 'least affected'; (ref)= reference category.

6.6 Summary

This chapter explored the baseline demographic characteristics of each latent state, presented the latent transition probabilities of moving between states, analysed stratified transition probabilities and examined whether the inclusion of four potentially confounding variables as a covariate altered the derived model.

Exploring the baseline demographic information revealed relationships with latent states. Notably, there were obvious differences in gender, living status, widespread pain, anxiety, depression, self-reported general health and sleep issues between the states. These predictors are further explored as potential factors of future state memberships in the next

chapter. The population studied here was comparable to other hand study populations, in that the individuals with more severe problems ('severely affected', 'high pain and poor gross function' and 'poor gross function') were more likely to be female and were generally older (Caspi et al., 2001; Zhang et al., 2002; Haara et al., 2003; Dahaghin 2005; Grotle et al., 2008a; Ghosh et al., 2014). It is important to note given the large sample size, even small differences become significant, and that the analyses presented in this chapter were unadjusted analyses (no additional characteristics were taken into account). Subsequent analyses (*Chapter 7*) explore adjusted relationships of baseline characteristics with latent states in more detail.

While large differences in latent transition probabilities between baseline to 3 years and 3 years to 6 years were not seen, in general a lower proportion of people transitioned into the 'high pain' or 'poor gross function' states at 6 years, and fewer remained within those latent states if they were classified as such at the previous time period. This implied that there was little transition into these two states, and individuals generally progressed from these states to the more severe conditions or improved into the 'least affected' state. The benefit of the third time period allowed the exploration of whether individuals' states fluctuated over time, or followed distinct paths through the various latent states. The most common example of individuals returning back to their baseline state at 6 years, was 'least affected' to 'high pain' to 'least affected' (n= 108). Therefore the 'high pain' without functional problems state may be an acute temporary phase for older adults.

Similarly, those with pain only ('high pain' state) appeared to have a higher probability of improving their hand symptoms compared to those with functional problems alone ('poor gross function' state), and those with hand pain and function problems ('high pain and poor gross function' state). This was reflected in the transition probabilities in *Table 6.3.1*, as the estimated transition probability into 'least affected' was higher for 'high pain'

compared to ‘poor gross function’ and ‘high pain and poor gross function’ during both time periods (τ= 0.42 for ‘high pain’, τ= 0.24 for ‘poor gross function’, τ= 0.21 for ‘high pain and poor gross function’ at baseline to 3 years, τ= 0.48 for ‘high pain’, τ= 0.28 for ‘poor gross function’, τ= 0.22 for ‘high pain and poor gross function’ at 3 years to 6 years). Also, those in the ‘high pain’ state were less likely to transition to the more severe states compared to those in the ‘poor gross function’ state at either time point (transition to ‘high pain and poor gross function’: τ= 0.15 for ‘high pain’, τ= 0.33 for ‘poor gross function’; transition to ‘severely affected’: τ= 0.02 for ‘high pain’, τ= 0.12 for ‘poor gross function’ at baseline to 3 years) (*Table 6.3.1*). This pattern was similar between 3 years and 6 years. Similarly, those in ‘high pain and poor gross function’ were more likely to transition to ‘severely affected’ compared to those from ‘high pain’ at each time period (τ= 0.21 vs. τ= 0.02 at baseline to 3 years, τ= 0.20 vs. τ= 0.05 at 3 years to 6 years, *Table 6.3.1*). In conclusion, this provided further indication that once individuals reported hand function problems, their prognosis was less favourable compared to those with hand pain alone.

It is reasonable to assume that a large proportion of the study population in problematic states had hand OA, especially considering the age and pain/ function issues of the individuals analysed. Hand OA (and OA in general) is not just characterised by long-term, persistent pain, but can also present with episodes (or ‘flares’) of increasing pain intensity and functional difficulties (Bellamy et al., 1997; Peat et al., 2001; Fautrel et al., 2005; Altman et al., 2009). These short term episodes of pain and function may also explain the long-term instability of the ‘high pain’ and ‘poor gross function’ states (in that the probabilities of individuals remaining in these states 3 years later are low in comparison to other states, ‘high pain’ τ= 0.38 (3 years) and τ= 0.26 (6 years), ‘poor gross function’ τ= 0.27 and τ= 0.35 (*Table 6.3.1*)).

Males with hand pain/ problems were less likely to remain in the 'severely affected' state at each transition, compared to the females. This result implies that in general, males had a higher chance of an improvement in their hand symptoms. In addition to this, individuals with no widespread pain were more likely to transition to, or stay in, the 'least affected' state, implying that they were also more likely to experience an improvement. Female gender and those with widespread pain have previously been shown to have a higher incidence of hand OA, and higher probability of progression (Grotle et al., 2008a; Nicholls et al., 2012).

The results in *section 6.4.4* indicated that those in the oldest age group were more likely to observe an improvement in their hand condition at a subsequent time point. However, it is important to note that the numbers for each of the problematic latent states for the oldest age group were small (all n<44) and inferences should be made with caution. Another potential reason for this increased probability was more severe participants were more likely to drop out of the study (*section 5.4, Table 5.4.1*), therefore, combined with the small sample size, these results may present an overly optimistic pattern for older people with hand pain/ problems.

The extension of transition probability stratification by characteristics such as gender is a useful and informative technique allowing the possibility of observing the different transitional patterns that sub-groups of individuals may take. By investigating the probabilities in this way, it was possible to assess whether pathways through the states were different for relevant sub-groups. Furthermore, carrying out this technique has highlighted that gender and widespread pain were associated with varying transitions to a future state. However, there were a few limitations to using this method as described below.

Firstly, the approach was quite computer intensive. With three time points and restrictions each model took numerous hours to compute; for example, when the factor has more than two levels (e.g. age group) this took 25-30 hours per simulation. In addition to this, each investigation required the computation of two models in order to test for a significant difference, therefore increasing the computation time. This was a practical reason for analysing only a small number of factors using this approach, focusing on those who were highlighted as being important predictors of the incidence or impact of hand OA by previous literature or the RUG.

Secondly, the information obtained from this approach was numerous (in an investigation split with 2 levels, such as gender, there were two sets of 125 pathways through baseline to 3 years to 6 years), which makes it difficult to interpret where key differences lie. Currently, there is no 'gold standard' on how much of a difference between transitional probabilities constitutes a meaningful difference (probabilities are also partially driven by state size). Therefore researchers need to develop their own subjective rules for assessing differences in the transition probabilities. This issue is exacerbated by the conflicting conclusions which may arise from BIC comparisons and LRT tests, requiring the need for an unrestricted model.

Additionally, a reasonably large sample (approximately 200, Hyatt and Collins, 1998; Collins and Lanza, 2010) is needed to compute a basic LTA. When the estimates produced in an LTA are stratified, or there are a high number of latent states, the sample needs to be increased to produce robust parameter estimates. For example, in the age analysis, the older age group estimates were based on just over 500 people. While this was generally sufficient, some issues with interpretation are evident due to the small numbers within some of the states, for example, there were just 17 people in the 'high pain' state from the

oldest age group at baseline. This may make the estimates imprecise as they were based on very small numbers if the frequencies were low.

Adjusting for a covariate in the latent state development (*section 6.5*) was an appropriate approach to investigating the accuracy of a LTA model, however it also incurs two drawbacks. The first drawback is that by including these covariates, the model becomes more complicated, and thus requires more time to compute. Resource time required alone should not prevent an examination of whether to include covariates or not. The second reason is interpretation of these models becomes more complex. For example, the model in *6.5.2* accounted for gender whilst developing the latent states, but gender was not being used to develop the latent states directly. It is possible to include more than one covariate into LTA modelling, however due to the increased computation time, and that the inclusion of one covariate had not been found to be essential, this aspect was not explored further in this project. In addition to this, including one covariate in the model complicates the interpretation, so including multiple covariates further complicates the model. Also, only a limited number of covariates can feasibly be analysed as separate models need to be derived (as presented in *section 6.5.2*).

Despite these drawbacks, it is advisable to explore (as in *section 6.5.2*) whether including covariates in the LTA model results in any changes in the estimated parameters (state proportions, item-response probabilities, transitional probabilities) or latent state membership. In this example, there was little benefit in adjusting for any of the covariates investigated within this latent state model, and therefore subsequent analysis does not include any adjusted factors, and uses the time-restricted base model that was presented in *section 6.3.1*.

Collectively, investigating different transitional patterns for factors of interest (gender, living status, age and widespread pain), and investigating these factors as potential covariates of latent state development has provided a good basis for the assessment of factors predictive of long term state, which is continued in the next chapter. The next chapter investigates how demographic characteristics, mood problems, general health and hand-specific factors may influence future hand state membership, and 6 year prognosis.

Chapter 7: **Exploring predictive factors of long-term hand state membership**

7.1 Introduction and objectives

The first part of this chapter describes an exploratory analysis of factors predictive of 3 and 6 year hand state membership. Factors examined include the demographic information, mood problems and general health that were briefly presented in the previous chapter, with the addition of hand factors and comorbidity measures. The second part of the chapter explores the predictive factors with improvement and deterioration over the 6 years follow-up of hand pain/ problems.

The specific objectives of the chapter were to:

i. Explore which factors predict 3 year and 6 year hand state membership;
ii. Explore which factors were related to a progression or an improvement in hand pain/ problems over 6 years.

7.2 Methods: Factors predicting 3 year and 6 year state membership

The first part of this chapter explores whether demographic characteristics, general health, lifestyle factors or specific hand complaints (excluding those that were used as indicators) were associated with membership in a particular latent hand state at 3 years, and to assess whether these factors were also associated with latent hand state at 6 years.

7.2.1 Inclusion criteria

The individuals included in all of the analyses in this chapter were those who were in one of the four 'problem' states at baseline, namely 'high pain', 'poor gross function', 'high pain and poor gross function' or 'severely affected' (therefore, this did not include the individuals classified in the 'least affected' state at baseline). Those in the 'least affected' state were excluded as these participants were considered unlikely to seek help from clinicians. Investigating which factors predict future hand states (and improvement, deterioration or stability) in individuals with hand pain/ problems increases the knowledge regarding the likelihood of future outcome of the condition (prognosis), which may be of practical clinical relevance for the people likely to consult with hand pain/ problems.

7.2.2 Potential factors of future state membership

All of the factors used to predict 3 or 6 year state membership were collected from the NorStOP baseline questionnaires, either from the HS or the RPS. The factors selected have been shown to be associated with future outcomes in painful MSK conditions (e.g. Kalichman and Hernández-Molina, 2010; Nicholls et al., 2012; Leung et al., 2013) or were considered by the research team to be potentially important as a predictive factor of long-term outcome.

For the ease of presentation only, baseline factors are presented in four groups: sociodemographic, health and lifestyle, hand-specific, and comorbidity factors. The factors are presented in *Table 7.2.1*. The first group of factors included demographic factors, most of which were reported in *Table 6.2.1*. The second group included other health and lifestyle factors (also shown in *Table 6.2.1*), with the four sleep problem items reduced into a binary variable representing individuals who responded to any of the four sleep questions

with problems on ‘most nights’, and those who did not (as per original scale, Jenkins et al., 1988).

The third group included factors that were taken from the more detailed hand questions contained in the RPS. As the analyses contained in this chapter excluded the least affected group at baseline and hence only included participants with hand pain/ problems at baseline; all of this population had received the RPS and had the opportunity to answer these specific hand questions.

The fourth group contained other specific health conditions (namely self-reported high blood pressure, diabetes, chest problems or heart problems) that were collected in the HS and were included for two main reasons. Firstly, the inclusion of these factors acted as a proxy for comorbidity, as they are four of the most frequent chronic health conditions (Ornstein et al., 2013; Roberts et al., 2014) that were available in the NorStOP database. Secondly, in previous work (Marshall et al., 2013; Visser 2013) a biological relationship between certain cardiovascular and metabolic factors, and the presence of hand pain/ problems was hypothesised. These four factors were combined to represent ‘any comorbidity’. This last assessment investigated whether the presence of at least one of these factors was associated with long-term hand state membership.

Table 7.2.1: Full list of factors with categories used in predictive analyses.

Group	Factor	Categories
1	Gender	Male, Female
	Age	50-64 years, 65-74 years, 75+ years
	Lived alone	Yes, No
	Employment status	Retired, Employed, Other (Ill/ Unemployed/ Housewife/ Other)
	Marital status	Married/ Cohabiting, Single (Separated/ Divorced/ Widowed/ Single)
	Social class	Higher managerial/ Professional, Intermediate, Routine/ Manual
2	ACR Widespread pain	Yes, No
	HADS	Anxiety and depression (two subscales, score range 0-21)
	BMI	Continuous scale
	Any sleep problems	Yes, No
	Frequency of GP visit	Often/ Very often, Occasionally/ Seldom/ Never
	SF-12 general health	Good/ Very good/ Excellent, Poor/ Fair
3	Previous hand injury	Yes, No
	Hand operation	Yes, No
	Excessive use (occupation or hobbies)	Yes, No
	Nodes	Yes, No
	Previous 12 month hand pain duration	<3 months, ≥ 3 months
	Bilateral hand pain	Yes, No
	Pain in two or more hand joints	No/ Few days, All/ Most/ Some days
	Impact of hand problem	Very well/ Well, Fair/ Poor/ Very poorly
	Self-report diagnosis of RA	Yes, No
	Frequency of medication use for hand symptoms	All/ Most days, Some/ Few/ No days
4	Comorbidity	Yes (Diabetes/ High blood pressure/ Heart problems/ Chest problems), No

Footnote: ACR= American College of Rheumatology (Wolfe et al., 1990); HADS= Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983); BMI= Body Mass Index; GP= General Practitioner; SF-12= Short-Form questionnaire (Ware et al., 1996); RA= Rheumatoid Arthritis.

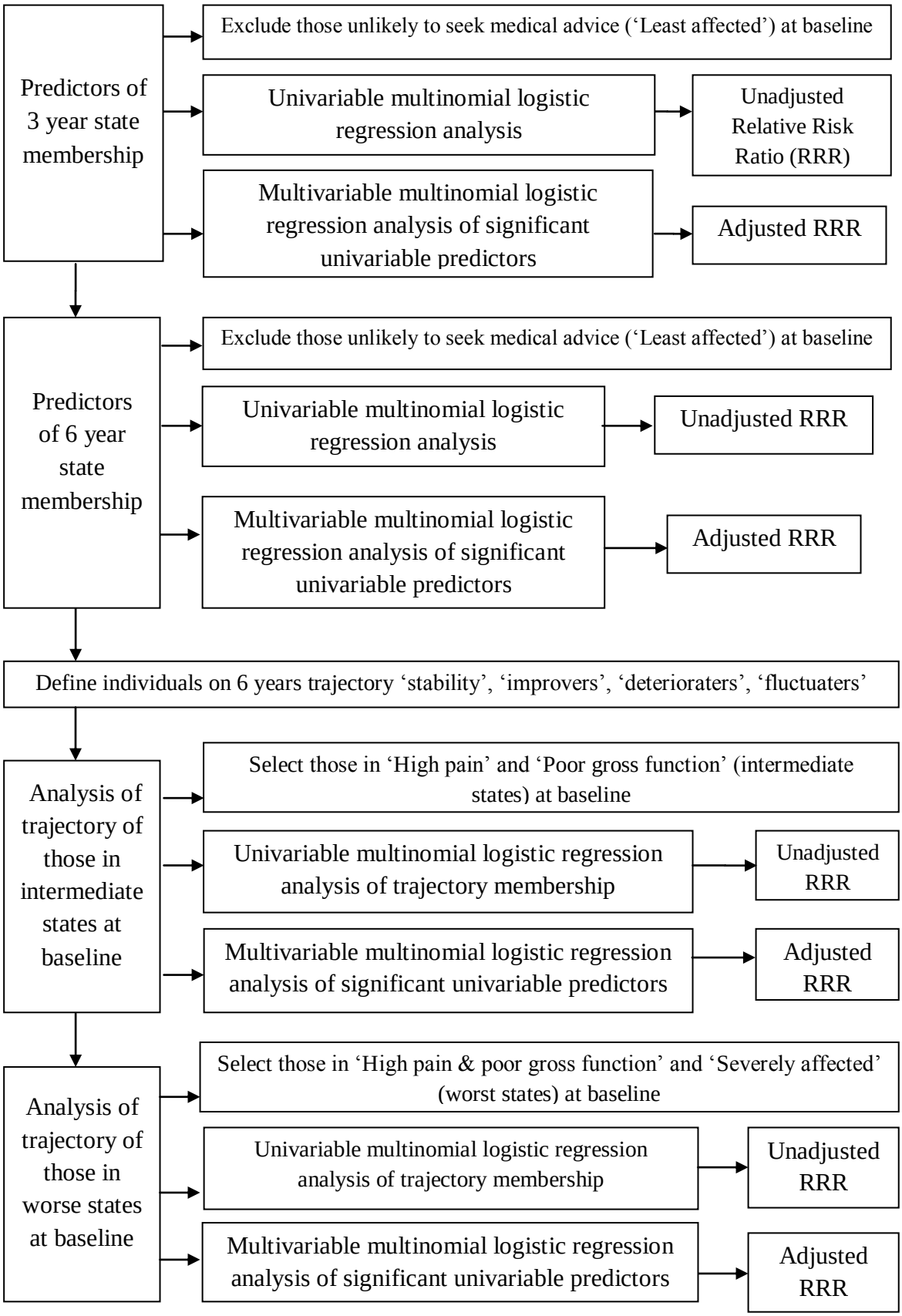
7.2.3 Analysis

Prior to the main analysis, the correlation between each of the potential predictive factors was assessed. As a correlation greater than 0.5 (or less than -0.5) is considered to be moderate to high correlation (Hinkle et al., 2003), baseline factors with correlations in this region were considered for removal. Multinomial logistic regression was used to investigate associations with state membership for all potential factors. This technique produces Relative Risk Ratios (RRR), which are presented with 95% confidence intervals.

All estimates were adjusted for baseline state, therefore exploring whether the association of each factor was associated with long term state independently of baseline state. For each assessment, 'least affected' at 3 years (and 6 years) was used as the reference. Therefore, considering that the population analysed excluded individuals in the 'least affected' state at baseline, the reference category reflected individuals who reported an improvement to 'least affected' in their hand symptoms at follow-up.

All analyses were adjusted for baseline state; hence the initial assessment of factors was not strictly a 'univariable' model. However, for ease of terminology this analysis is referred to as such. After the univariable assessment of each of the factors, the factors that were significantly associated with latent state membership at 3 years (and 6 years) were taken forward into a full multivariable multinomial regression model (second stage of analysis). Factors taken forward were those with a p-value<0.05 for association of membership in at least one of the problem states compared to 'least affected' in the initial (adjustment for baseline state only) analysis. After this stage, no further factors were removed. For all of the analyses in this chapter, statistical significance was regarded as p-value<0.05. STATA version 13.1 was used for the analysis in this chapter (StataCorp, 2013). The analyses in this chapter is summarised in *Figure 7.2.1*.

Figure 7.2.1: Flow diagram of the stages of analysis performed in Chapter 7.



Footnote: Left box for each denotes the overall title for each stage, the first flow of each stage specifies the population for that analysis, with the remaining flows highlighting the statistical technique and the resulting estimates that were produced.

7.3 Results: Factors predicting 3 year state membership (Univariable associations)

7.3.1 Results - Sociodemographic factors

The correlation between marital status and living status was 0.85, therefore, due to marital status having a weaker unadjusted association with 3 year hand state, this factor was removed. In addition to this, the correlation between employment status and age was -0.52. However as this correlation was on the borderline of the pre-specified level and the factors may have had a different impact on hand problems, it was agreed with supervisors that both of these factors would remain as potential factors in the model.

The results from the first set of factors, adjusted only for baseline state, are presented in *Table 7.3.1* (the frequency and proportion of individuals in each state for each factor are shown in *Appendix F*). Similar to results found in *Chapter 6*, gender had an impact on 3 year latent state membership, with males significantly less likely to be in the 'poor gross function' state (RRR= 0.59, 95%CI:(0.36, 0.97)), and less likely to be in the 'severely affected' state (albeit not significant at the 5% level RRR= 0.70 (0.47, 1.05)), than being in the 'least affected' state compared to females. The oldest age groups were less likely to be in the worst states compared to the youngest group (RRR's range 0.09 to 0.90), and individuals who were not retired, were more likely to be in the 'high pain' state at 3 years (employed RRR= 2.79 (1.61, 4.83), others RRR= 2.02 (1.02, 4.00)).

Social class and living alone had little impact on 3 year state membership after adjustment for baseline state. Therefore, the factors that were taken forward to the final full multivariable model were gender, age and employment status.

Table 7.3.1: Results from group one: Associations (RRR (95% CI)) of demographic factors with latent state at 3 years, adjusting for baseline state only.

Baseline factor		LA	High pain	Poor gross function	HPPGF	Severely affected
Gender	Male	1.00	1.38	0.59	0.92	0.70
			(0.85,2.25)	(0.36,0.97)	(0.65,1.31)	(0.47,1.05)
Age (years)	50-64	1.00	1.00	1.00	1.00	1.00
	65-74	1.00	0.68	0.65	0.64	0.80
			(0.40,1.16)	(0.41,1.04)	(0.45,0.91)	(0.54,1.18)
	75+	1.00	0.09	0.62	0.41	0.90
			(0.02,0.41)	(0.32,1.24)	(0.23,0.71)	(0.51,1.59)
Lived alone	Yes	1.00	0.66	0.89	0.82	0.93
			(0.33,1.31)	(0.53,1.50)	(0.55,1.23)	(0.60,1.43)
Employment status	Retired	1.00	1.00	1.00	1.00	1.00
	Employed	1.00	2.79	1.43	1.80	0.86
			(1.61,4.83)	(0.84,2.42)	(1.20,2.72)	(0.52,1.44)
	Other	1.00	2.02	0.92	1.39	1.09
			(1.02,4.00)	(0.53,1.61)	(0.92,2.09)	(0.71,1.70)
Social class	Higher managerial/ Professional	1.00	1.00	1.00	1.00	1.00
	Intermediate	1.00	0.80	0.55	0.66	0.73
			(0.41,1.53)	(0.29,1.06)	(0.41,1.08)	(0.41,1.29)
	Routine/ Manual	1.00	0.62	0.87	0.80	0.76
			(0.35,1.11)	(0.50,1.52)	(0.52,1.24)	(0.46,1.27)

Footnote: RRR= Relative Risk Ratio; CI= Confidence Interval; LA= 'least affected'; HPPGF= 'high pain and poor gross function'.

7.3.2 Results - Health and life style factors

The depression and anxiety scores were highly correlated (0.65) and therefore, as depression score was more strongly associated with 3 year (and 6 year) hand state, anxiety score was removed from the analysis.

Several general health and lifestyle factors (displayed in *Table 7.3.2*) were associated with 3 year latent states after adjustment for baseline state. Firstly, those classified as having widespread pain at baseline were significantly more likely to be in a problem hand state compared to being in the 'least affected' state (all RRR>1.40), with an increased trend towards the more severe states. Higher scores on the depression scale appeared to have a modest association with hand states with pain elements, specifically 'high pain' (RRR for each additional unit of depression score= 1.05 (0.98, 1.13), *p*-value= 0.17) and 'severely affected' (RRR= 1.04 (0.99, 1.09) *p*-value= 0.14), while a lower score was associated with 'poor gross function' (RRR= 0.94 (0.88, 1.01, *p*-value= 0.08)). Higher BMI was significantly associated only with 'high pain and poor gross function' RRR= 1.03 (1.00, 1.07 per unit increase in BMI), *p*-value= 0.03). It is worth noting that the depression scale and BMI were measured on a continuous scale, so these estimates reflect a per point increase (normality for BMI was acceptable, depression score was negatively skewed, however, due to Central Limit Theorem and a large sample size (Durrett, 2004) this factor was not modified and a linear relationship was presumed). Given this, it was agreed that both of these two factors were to be taken forward to the multivariable model as associations with BMI was significant for 'high pain and poor gross function' and depression score showed borderline significance with several hand states.

Reporting a sleep problem at baseline was associated with membership in 'high pain and poor gross function' (RRR= 1.36 (0.98, 1.89)) and significantly associated with 'severely affected' (RRR= 1.72 (1.19, 2.48)) at 3 years. Also, if a participant regarded their overall health to be 'poor' or 'fair', they were over twice as likely to be in the 'severely affected' state at 3 years (RRR= 2.35 (1.62, 3.43)), compared to being in the 'least affected' state. Self-reported GP visits had no significant association with 3 year hand state.

In addition to BMI and the depression scale, widespread pain, any sleep problems, and self-perceived health status were taken through to the multivariable model.

Table 7.3.2: Results from group two: Associations (RRR (95% CI)) of general health factors with latent state at 3 years, adjusting for baseline state.

Baseline factor		LA	High pain	Poor gross function	High pain & poor gross function	Severely affected
ACR Widespread pain		1.00	1.56	1.40	2.06	2.03
			(0.97,2.52)	(0.91,2.15)	(1.48,2.86)	(1.40,2.94)
HADS	Depression	1.00	1.05	0.94	1.00	1.04
			(0.98,1.13)	(0.88,1.01)	(0.96,1.05)	(0.99,1.09)
BMI		1.00	1.04	1.00	1.03	1.02
			(0.98,1.09)	(0.96,1.05)	(1.00,1.07)	(0.98,1.06)
Any sleep problems		1.00	1.34	1.22	1.36	1.72
			(0.82,2.17)	(0.80,1.88)	(0.98,1.89)	(1.19,2.48)
Frequency of GP visit	Occasionally/ Seldom/ Never	1.00	1.00	1.00	1.00	1.00
	Often/ Very often	1.00	0.64	0.68	0.99	1.12
			(0.33,1.22)	(0.40,1.15)	(0.68,1.45)	(0.74,1.69)
SF-12 general health	Good/ Very good/ Excellent	1.00	1.00	1.00	1.00	1.00
	Poor/ Fair	1.00	0.88	0.64	1.02	2.35
			(0.52,1.52)	(0.40,1.03)	(0.72,1.44)	(1.62,3.43)

Footnote: RRR= Relative Risk Ratio; CI= Confidence Interval; LA= 'least affected'; ACR= American College of Rheumatology (Wolfe et al., 1990); HADS= Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983); BMI= Body Mass Index; GP= General Practitioner; SF-12= Short Form questionnaire (Ware et al., 1996).

7.3.3 Results - Hand-specific factors

The associations of hand-specific factors with 3 year state membership displayed some unexpected results (*Table 7.3.3*). 'Previous hand injury', 'hand operation' and 'excessive use' (either in occupation or hobbies), were not significantly associated with membership

in any of the problem states at 3 years, with no apparent trend through the severity of states (no pattern of increasing/ decreasing RRR over the states). A potential explanation for this is that those with these hand-specific factors were more likely to be in 'severely affected' or 'high pain and poor gross function' at baseline, therefore individuals were already in the more severe states (*Appendix F*). The presence of finger nodes displayed a strong association with membership in a problem state at 3 years (all RRR>1.51 and most significant at 5% level).

The remaining hand-specific factors all had a significant association with membership in the majority of the problem states, with the highest association being for those with bilateral hand pain (all RRR>1.53), with over twice the probability of belonging to the 'severely affected' state at 3 years than the 'least affected' state (RRR= 2.37 (1.58, 3.57)). Similarly, those who reported pain in two or more hand joints on all, most or some days were over twice as likely to be in 'high pain and poor gross function' or 'severely affected' at 3 years (RRR= 2.02 (1.42, 2.88) and RRR= 2.16 (1.42, 3.29)). Participants with chronic hand pain (≥3 months of hand pain in last 12 months) were more likely to belong to a more severe group (especially those with a strong pain component, (not 'poor gross function')). Those who classified themselves in the 'fair', 'poor' or 'very poor' category instead of 'very well' or 'well' with regard to the impact of their hand condition, were significantly more likely to be in the 'severely affected' state at 3 years than in the 'least affected'(RRR= 1.55 (1.04, 2.31). Self-reported diagnosis of RA had an association with membership in 'high pain' (RRR= 1.80 (0.99, 3.25), *p*-value= 0.052) and as such was included in the multivariable model, along with frequently taking medication for hand symptoms on most or all days ('high pain and poor gross function', RRR= 2.23 (1.45, 3.42), 'severely affected', RRR= 2.24 (1.44, 3.48)).

Table 7.3.3: Results from group three: Associations (RRR (95% CI)) of hand-specific factors with latent state at 3 years, adjusting for baseline state.

Baseline factor		LA	High pain	Poor gross function	High pain & poor gross function	Severely affected
Previous hand injury	Yes	1.00	0.97	0.97	1.12	1.19
			(0.56,1.67)	(0.60,1.55)	(0.78,1.60)	(0.80,1.76)
Hand operation	Yes	1.00	1.03	0.95	1.01	0.89
			(0.53,2.00)	(0.54,1.68)	(0.66,1.56)	(0.55,1.43)
Excessive use	Yes	1.00	0.87	0.70	1.35	1.02
			(0.46,1.65)	(0.41,1.22)	(0.83,2.17)	(0.60,1.71)
Nodes	Yes	1.00	1.51	1.58	1.48	2.17
			(0.91,2.50)	(1.02,2.44)	(1.05,2.07)	(1.49,3.15)
Previous 12 month hand pain duration	<3 months	1.00	1.00	1.00	1.00	1.00
	≥3 months	1.00	1.56	1.41	1.60	2.11
			(0.94,2.58)	(0.89,2.25)	(1.12,2.29)	(1.37,3.27)
Bilateral hand pain	Yes	1.00	1.53	1.67	1.88	2.37
			(0.93,2.52)	(1.05,2.67)	(1.32,2.67)	(1.58,3.57)
Pain in two or more hand joints	No/ Few days	1.00	1.00	1.00	1.00	1.00
	All/ Most/ Some days	1.00	1.55 (0.95,2.53)	1.40 (0.88,2.21)	2.02 (1.42,2.88)	2.16 (1.42,3.29)
Impact of hand problem	Very well/ Well	1.00	1.00	1.00	1.00	1.00
	Fair/ Poor/ Very poorly	1.00	1.49	0.89	1.07	1.55
			(0.78,2.84)	(0.53,1.49)	(0.73,1.58)	(1.04,2.31)
RA		1.00	1.80	1.01	1.21	1.52
			(0.99,3.25)	(0.55,1.87)	(0.77,1.88)	(0.94,2.46)
Frequency of medication use	Some/ Few/ No days	1.00	1.00	1.00	1.00	1.00
	All/ Most days	1.00	0.81	1.07	2.23	2.24
			(0.36,1.80)	(0.58,1.98)	(1.45,3.42)	(1.44,3.48)

Footnote: RRR= Relative Risk Ratio; CI= Confidence Interval; LA= 'least affected'; RA= Rheumatoid Arthritis.

The factors to be included in the multivariable model from this group were presence of nodes, chronic hand pain duration, bilateral hand pain, pain in two or more hand joints, perceived general health (SF-12), self-reported diagnosis of RA and frequent medication use for hand symptoms.

7.3.4 Results - Comorbidity factor

Comorbidity (diabetes, high blood pressure, heart or chest problems) was not associated with membership in a pain or function hand state at 3 years (*Table 7.3.4*). Therefore, it did not appear that having a common comorbid condition increased (or decreased) an individuals' likelihood of 3 year state membership, and this factor was not taken forward to be included in the multivariable model.

Table 7.3.4: Results from group four: Univariable associations (RRR (95% CI)) of additional health factors with latent state at 3 years, adjusting for baseline state.

Baseline factor	LA	High pain	Poor gross function	High pain & poor gross function	Severely affected
Any comorbidity	1.00	0.89 (0.55,1.43)	0.94 (0.61,1.44)	1.12 (0.81,1.56)	1.03 (0.71,1.49)

Footnote: RRR= Relative Risk Ratio; CI= Confidence Interval; LA= 'least affected'.

7.3.5 Factors predicting 3 year state membership (Multivariable)

The factors taken forward from the four groups were included in a multivariable model and the results are displayed in *Table 7.3.5*. Of the 1,279 participants, 1,042 (81%) had complete data and were included in the multivariable model. The following baseline characteristics were significantly associated with membership of at least one of the problem hand states at 3 years: age, widespread pain, symptoms of depression, sleep problems, self-perceived health status, nodes, chronic hand pain duration, bilateral hand pain and frequent medication use for hand symptoms.

Table 7.3.5: Results (RRR (95% CI)) from the multivariable model of factors of latent state at 3 years, adjusting for baseline state.

Baseline factor (n= 1,042 at 3 years)		LA (n= 224)	HP (n= 99)	PGF (n= 111)	HPPGF (n= 301)	SA (n= 307)
Gender	Male	1.00	1.55 (0.86,2.78)	0.88 (0.48,1.63)	1.24 (0.80,1.93)	0.93 (0.56,1.54)
Age (years)	50-64	1.00	1.00	1.00	1.00	1.00
	65-74	1.00	0.86 (0.37,1.97)	**0.47 (0.23,0.94)**	**0.53 (0.31,0.91)**	**0.50 (0.27,0.94)**
	75+	1.00	**0.07 (0.01,0.64)**	0.72 (0.29,1.79)	**0.44 (0.20,0.96)**	0.94 (0.41,2.14)
Employment status	Retired	1.00	1.00	1.00	1.00	1.00
	Employed	1.00	1.72 (0.77,3.86)	0.91 (0.43,1.94)	1.06 (0.58,1.92)	0.76 (0.37,1.58)
	Other	1.00	1.20 (0.48,3.03)	0.73 (0.34,1.55)	0.79 (0.44,1.42)	0.70 (0.36,1.36)
ACR Widespread pain		1.00	1.64 (0.91,2.94)	1.37 (0.82,2.31)	**1.95 (1.31,2.91)**	1.46 (0.92,2.30)
HADS	Depression	1.00	1.04 (0.93,1.15)	**0.91 (0.83,0.99)**	0.95 (0.89,1.02)	0.97 (0.90,1.04)
BMI		1.00	1.03 (0.97,1.09)	1.00 (0.95,1.05)	1.02 (0.99,1.06)	0.99 (0.95,1.04)
Any sleep problems		1.00	1.29 (0.70,2.37)	**1.91 (1.11,3.27)**	**1.62 (1.06,2.46)**	**2.05 (1.27,3.31)**
SF-12 general health	G/ VG/ Excel	1.00	1.00	1.00	1.00	1.00
	Poor/ Fair	1.00	0.69 (0.33,1.42)	1.03 (0.55,1.95)	1.00 (0.63,1.61)	**2.58 (1.54,4.31)**
Nodes	Yes	1.00	1.58 (0.88,2.85)	1.30 (0.77,2.20)	1.38 (0.91,2.08)	**2.27 (1.42,3.64)**
Previous 12 month hand pain duration	<3 months	1.00	1.00	1.00	1.00	1.00
	≥3 months	1.00	1.71 (0.93,3.15)	1.58 (0.91,2.74)	**1.59 (1.03,2.44)**	**2.13 (1.25,3.65)**
Bilateral hand pain		1.00	1.44 (0.77,3.15)	1.37 (0.79,2.39)	1.24 (0.81,1.89)	**1.71 (1.03,2.84)**
Pain in two or more hand joints	No/ Few days	1.00	1.00	1.00	1.00	1.00
	All/ Most/ Some days	1.00	1.12 (0.60,2.07)	0.89 (0.50,1.57)	1.22 (0.78,1.90)	1.04 (0.61,1.76)
Impact of hand problem	Very well/ Well	1.00	1.00	1.00	1.00	1.00
	Fair/ Poor/ Very poorly	1.00	0.89 (0.39,2.04)	0.77 (0.40,1.47)	0.86 (0.53,1.39)	1.17 (0.70,1.94)
RA		1.00	1.95 (0.96,3.97)	1.41 (0.71,2.82)	1.38 (0.81,2.35)	1.73 (0.96,3.14)
Frequency of medication use	Some/ Few/ No days	1.00	1.00	1.00	1.00	1.00
	All/ Most days	1.00	0.60 (0.22,1.61)	0.99 (0.49,2.00)	**1.81 (1.09,3.00)**	1.31 (0.77,2.25)
Time 1 state	HP	1.00	1.00	1.00	1.00	1.00
	PGF	1.00	**0.13 (0.05,0.34)**	**28.50 (9.37,86.67)**	**7.31 (3.88,13.75)**	**16.09 (4.50,57.57)**

Baseline factor (n= 1,042 at 3 years)	LA (n= 224)	HP (n= 99)	PGF (n= 111)	HPPGF (n= 301)	SA (n= 307)
HPPGF	1.00	**0.08 (0.03,0.21)**	**9.16 (2.96,28.32)**	**6.75 (3.73,12.21)**	**16.63 (4.84,57.20)**
SA	1.00	**0.03 (0.00,0.22)**	**10.39 (2.94,36.68)**	**2.19 (1.01,4.75)**	**88.78 (24.86,317.04)**

Footnote: RRR= Relative Risk Ratio; CI= Confidence Interval; LA= 'least affected'; HP= 'high pain'; PGF= 'poor gross function'; HPPGF= 'high pain and poor gross function'; SA= 'severely affected'; ACR= American College of Rheumatology (Wolfe et al., 1990); HADS= Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983); BMI= Body Mass Index; SF-12= Short Form questionnaire (Ware et al., 1996); G= Good; VG= Very good; Excel= Excellent; RA= Rheumatoid Arthritis; Bold entries represent significant at 5% after adjustment for other factors in table.

Firstly, individuals aged 65-74 (the 'newly retired' group) were significantly less likely to belong in any of the functional problem states ('poor gross function' RRR= 0.47 (0.23, 0.94), 'high pain and poor gross function' RRR= 0.53 (0.31, 0.91), or 'severely affected' RRR= 0.50 (0.27, 0.94)) than to have recovered (i.e. moved to 'least affected' state). In addition to this, those in the oldest age group (aged 75+) were significantly less likely to be in the 'high pain' state at 3 years than in the 'least affected' state (RRR= 0.07 (0.01, 0.64)).

Although not all estimates were significant, the presence of widespread pain increased the likelihood of being in one of the problem states (i.e. not 'least affected') at 3 years (all RRR≥ 1.37), as did presence of bilateral hand pain (all RRR≥ 1.24). The presence of sleep problems had a strong significant association with membership in the functional problem hand states, with membership in 'severely affected' twice as likely compared to being in the 'least affected' state (RRR= 2.05 (1.27, 3.31)). In addition, chronic hand pain duration had a clear trend with increased membership in the more severe groups (from 'high pain' RRR= 1.71 (0.93, 3.15), *p*-value= 0.085 to 'severely affected' RRR= 2.13 (1.25, 3.65), *p*-value= 0.006).

The presence of finger nodes appeared to be strongly associated with 3 year membership in the 'severely affected' state at 3 years (RRR= 2.27 (1.42, 3.64)). Participants who reported frequent medication use for their hand symptoms (on 'most' or 'all' days) were more likely

to be in the more severe states ('high pain and poor gross function' RRR= 1.81 (1.09, 3.00), 'severely affected' RRR= 1.31 (0.77, 2.25)).

A clear finding from the estimates was that baseline state membership was a strong factor of hand state at 3 years, especially for 'severely affected' (RRR= 88.78 (24.86, 317.04) compared to being in 'high pain' state at baseline). While the number of observations may be small, thus a wide confidence interval, this finding aligns with that from *Chapter 6* indicating that people in the 'severely affected' state were more likely to remain in that state at 3 years. In addition to this, the estimates support the previous findings that once an individual had hand function difficulties, they were unlikely to improve, whilst those in 'high pain' were more likely to see an improvement (RRRs at the bottom of *Table 7.3.5*).

7.3.6 Factors predicting 6 year state membership (Multivariable)

For brevity, the result of the univariable analysis for 6 year state membership of the four groups of factors is presented in *Appendix F*. The resulting multivariable model of all the factors that were taken forward after univariable analysis is presented in *Table 7.3.6*. Firstly, the factors included in the 6 years multivariable model were almost identical to those in the 3 years multivariable model. The only exceptions were depression score, which was no longer included in the 6 years predictive model, and social class which was included in the 6 years model (but not previously in the 3 years model). The 6 years model was similar to the 3 years model but the strength of some of the associations were weaker.

Table 7.3.6: Results (RRR (95% CI)) from the multivariable model of factors of latent state at 6 years, adjusting for baseline state.

Baseline factor (n= 999 at 6 years)		LA (n= 264)	HP (n= 72)	PGF (n= 94)	HPPGF (n= 268)	SA (n= 301)
Gender	Male	1.00	1.81 (0.99,3.31)	0.79 (0.44,1.42)	0.69 (0.46,1.03)	**0.52 (0.33,0.80)**
Age (years)	50-64	1.00	1.00	1.00	1.00	1.00
	65-74	1.00	0.71 (0.30,1.68)	0.89 (0.44,1.80)	0.77 (0.46,1.30)	0.75 (0.43,1.30)
	75+	1.00	0.23 (0.05,1.15)	**0.34 (0.11,0.99)**	**0.46 (0.22,0.95)**	0.81 (0.39,1.67)
Employment status	Retired	1.00	1.00	1.00	1.00	1.00
	Employed	1.00	1.61 (0.70,3.67)	0.88 (0.41,1.89)	1.03 (0.60,1.78)	0.71 (0.38,1.32)
	Other	1.00	1.23 (0.48,3.10)	0.65 (0.29,1.46)	1.10 (0.63,1.91)	0.80 (0.44,1.44)
Social class	Higher man/ Professional	1.00	1.00	1.00	1.00	1.00
	Intermediate	1.00	0.88 (0.41,1.89)	**0.43 (0.21,0.87)**	0.76 (0.46,1.27)	0.72 (0.41,1.28)
	Routine/ Manual	1.00	0.66 (0.33,1.34)	**0.53 (0.29,0.98)**	0.69 (0.43,1.11)	0.87 (0.52,1.45)
ACR Widespread pain		1.00	1.35 (0.75,2.42)	1.15 (0.68,1.96)	1.11 (0.76,1.62)	1.27 (0.84,1.91)
BMI		1.00	1.04 (0.99,1.09)	0.97 (0.92,1.02)	0.99 (0.96,1.03)	0.99 (0.96,1.03)
Any sleep problems		1.00	0.69 (0.38,1.28)	1.41 (0.83,2.37)	**1.62 (1.11,2.36)**	**1.50 (1.00,2.25)**
SF-12 general health	G/ VG/ Excel	1.00	1.00	1.00	1.00	1.00
	Poor/ Fair	1.00	0.56 (0.28,1.12)	0.73 (0.40,1.34)	0.69 (0.45,1.05)	1.37 (0.88,2.13)
Nodes		1.00	1.76 (0.96,3.12)	1.70 (1.00,2.90)	**1.53 (1.04,2.23)**	**2.08 (1.38,3.14)**
Previous 12 month hand pain duration	<3 months	1.00	1.00	1.00	1.00	1.00
	≥3 months	1.00	1.26 (0.67,2.38)	1.18 (0.67,2.07)	**1.58 (1.04,2.38)**	1.31 (0.83,2.09)
Bilateral hand pain		1.00	0.72 (0.39,1.33)	1.11 (0.63,1.95)	**1.56 (1.03,2.36)**	**1.67 (1.06,2.63)**
Pain in two or more hand joints	No/ Few days	1.00	1.00	1.00	1.00	1.00
	All/ Most/ Some days	1.00	1.19 (0.63,2.24)	0.86 (0.48,1.53)	1.06 (0.69,1.61)	1.09 (0.68,1.75)
Impact of hand problem	Very well/ Well	1.00	1.00	1.00	1.00	1.00
	Fair/ Poor/ Very poorly	1.00	1.88 (0.89,3.98)	**2.44 (1.32,4.51)**	1.14 (0.71,1.83)	1.35 (0.84,2.16)
Frequency of medication use	Some/ Few/ No days	1.00	1.00	1.00	1.00	1.00
	All/ Most days	1.00	0.83 (0.34,2.00)	1.14 (0.59,2.23)	1.11 (0.69,1.77)	1.11 (0.70,1.78)

Baseline factor (n= 999 at 6 years)		LA (n= 264)	HP (n= 72)	PGF (n= 94)	HPPGF (n= 268)	SA (n= 301)
Time 1 state	HP	1.00	1.00	1.00	1.00	1.00
	PGF	1.00	**0.18 (0.06,0.49)**	**9.70 (3.79,24.81)**	**2.05 (1.20,3.49)**	**3.39 (1.57,7.34)**
	HPPGF	1.00	**0.46 (0.23,0.94)**	**4.71 (1.78,12.43)**	**2.65 (1.59,4.42)**	**5.26 (2.52,10.96)**
	SA	1.00	**0.15 (0.04,0.57)**	2.06 (0.61,7.00)	1.36 (0.69,2.66)	**17.32 (7.87,38.11)**

Footnote: RRR= Relative Risk Ratio; CI= Confidence Interval; LA= 'least affected'; HP= 'high pain'; PGF= 'poor gross function'; HPPGF= 'high pain and poor gross function'; SA= 'severely affected'; man= managerial; ACR= American College of Rheumatology (Wolfe et al., 1990); BMI= Body Mass Index; SF-12= Short Form questionnaire (Ware et al., 1996); G= Good; VG= Very good; Excel= Excellent; Bold entries represent significant at 5% after adjustment for other factors in table.

The baseline factors that remained highly associated with 6 years state were sleep problems with functional problem states (RRR> 1.41, except 'high pain'), presence of nodes with all problem states (all RRR> 1.53) and chronic hand pain duration (more than 3 months) with all states including hand pain (RRR> 1.26). In addition to this, individuals with bilateral hand pain were more likely to be a member of the two worst hand states: 'high pain and poor gross function' (RRR= 1.56 (1.03, 2.36)) and 'severely affected' (RRR= 1.67 (1.06, 2.63)). Similar to the 3 years multivariable model, the strongest factor at 6 years was baseline hand state.

7.4 Methods: Factors predicting 6 year trajectories

The purpose of the second half of this chapter was to investigate factors of the 6 year course of hand pain/ problems over the follow-up period. Carrying out this analysis, in addition to that already performed in this chapter, provides an alternative approach of identifying the outcomes of individuals with hand pain/ problems, and identifying factors that are associated with those outcomes. The two approaches are subtly different. The first half of this chapter investigated which factors were associated with membership in follow-up states; the second half of the chapter investigates which factors were associated with

membership in an overall improving or declining trajectory of hand pain/ problems over the three different time points.

For this analysis the same study population was used, but split into two groups: those who were in the 'high pain' and 'poor gross function' (medium/ intermediate) states at baseline, and those in the 'high pain and poor gross function' and 'severely affected' (more severe/ worst) states at baseline. In these two respective groups the purpose was to determine the frequency of each 6 year course/ trajectory, whether they improved, deteriorated, fluctuated or maintained their hand states over the 6 years, and also to explore which factors were associated with membership in each trajectory.

The trajectories were developed from the latent states observed at baseline, 3 years and 6 years. However, as previous results have shown similar characteristics between 'high pain and poor gross function' and 'severely affected', these two states were combined to represent the 'worst states'. It is difficult to determine which states of 'high pain' and 'poor gross function' has a greater impact on individuals, therefore these states were also combined to represent 'intermediate' states.

7.4.1 Analysis

The same population used in *section 7.3* was used in this analysis, therefore excluding participants in 'least affected' state at baseline. The first trajectory outcome represented 'stability' over the 6 years. Individuals classified in this trajectory were either in the intermediate group at all three time points or the worst group at all three time points. The second trajectory outcome, labelled 'improvers', represented the individuals that improved over the course of the 6 years compared to baseline. Conversely, the third trajectory, labelled 'deterioraters', contained individuals who were in a more severe group at 6 years compared to baseline. Finally, the fourth trajectory represented individuals whose group

membership fluctuated over the 6 years, and thus were labelled ‘fluctuaters’. Individuals in this trajectory may have experienced an improvement between baseline and 3 years, only to deteriorate again between 3 years and 6 years (or vice versa).

The baseline characteristics of the individuals in each of these trajectories were assessed, followed by a multinomial logistic process to determine the significant univariable factors to explore in a multivariable model (similar to *section 7.3*). However, these investigations were stratified into individuals who were in the ‘intermediate’ group (‘high pain’ and ‘poor gross function’) at baseline and individuals who were in the ‘worst’ group (‘high pain and poor gross function’ and ‘severely affected’) at baseline. Those who were in the ‘intermediate’ group at baseline could improve, deteriorate, fluctuate or remain stable, while those in the ‘worst’ group could improve, fluctuate or remain stable. The development of the model was similar to the stages described in *section 7.3* and for brevity only the multivariable (adjusted) models are presented for each of the two analyses (the univariable results are reported in *Appendix G*) with the key significant factors highlighted.

7.5 Results: Factors predicting 6 year trajectories

7.5.1 Intermediate baseline group - Baseline characteristics

Of the participants who were in one of the intermediate hand states at baseline (‘high pain’ or ‘poor gross function’), there were four potential trajectories in which they could have been classified, indicated in *Table 7.5.1*. 169 individuals (32% of analysis population) improved their hand state over the 6 year period, 183 deteriorated (34%), 101 fluctuated (19%) and 78 remained stable (15%).

Table 7.5.1: Baseline characteristics of participants in each of the trajectories for those in the intermediate baseline groups (n (%) unless stated).

Baseline factor (n= 531)		Improvers (n= 169) (31.8%)	Stable (n= 78) (14.7%)	Deteriorators (n= 183) (34.5%)	Fluctuaters (n= 101) (19.0%)
Age (years)	50-64	100 (59.2%)	51 (65.4%)	106 (57.9%)	50 (49.5%)
	65-74	51 (30.2%)	24 (30.8%)	57 (31.2%)	38 (37.6%)
	75+	18 (10.7%)	3 (3.9%)	20 (10.9%)	13 (12.9%)
Gender	Male	76 (45.0%)	33 (42.3%)	47 (25.7%)	38 (37.6%)
Lived alone	Yes	28 (17.6%)	12 (15.8%)	37 (20.8%)	12 (12.0%)
Marital status	Married/ Cohabiting	130 (77.4%)	64 (82.1%)	133 (73.1%)	81 (80.2%)
	Single	38 (22.6%)	14 (18.0%)	49 (26.9%)	20 (19.8%)
Employment status	Retired	80 (48.5%)	35 (45.5%)	94 (52.2%)	50 (51.0%)
	Employed	56 (33.9%)	31 (40.3%)	54 (30.0%)	29 (29.6%)
	Other	29 (17.6%)	11 (14.3%)	32 (17.8%)	19 (19.4%)
Social class	Higher managerial/ Professional	43 (26.5%)	25 (32.5%)	44 (25.4%)	29 (29.6%)
	Intermediate	49 (30.3%)	20 (26.0%)	43 (24.9%)	21 (21.4%)
	Routine/ Manual	70 (43.2%)	32 (41.6%)	86 (49.7%)	48 (49.0%)
HADS Depression (mean (SD))		3.96 (3.0)	3.99 (2.5)	4.65 (3.4)	4.56 (3.5)
ACR Widespread pain		67 (39.6%)	36 (46.2%)	93 (50.8%)	50 (49.5%)
BMI (mean (SD))		26.7 (5.2)	27.1 (5.1)	26.6 (4.4)	28.3 (4.5)
Any sleep problems		66 (39.1%)	30 (38.5%)	98 (54.1%)	38 (37.6%)
Previous hand injury		43 (26.4%)	26 (33.8%)	50 (28.4%)	34 (35.1%)
Hand operation		23 (14.2%)	15 (19.5%)	29 (16.5%)	17 (17.5%)
Excessive use		135 (80.4%)	63 (80.8%)	150 (82.9%)	83 (83.0%)
Nodes		58 (36.5%)	40 (52.0%)	96 (54.6%)	40 (40.8%)
Previous 12 month hand pain duration	<3 months	75 (46.6%)	30 (39.5%)	58 (32.8%)	43 (44.3%)
	≥3 months	86 (53.4%)	46 (60.5%)	119 (67.2%)	54 (55.7%)
Bilateral hand pain		93 (59.2%)	48 (63.2%)	132 (75.4%)	56 (58.3%)
Pain in two or more hand joints	No/ Few days	94 (58.0%)	38 (50.7%)	84 (46.2%)	56 (57.1%)
	All/ Most/ Some days	68 (42.0%)	37 (49.3%)	98 (53.9%)	42 (42.9%)
Impact of hand problem	Very well/ Well	144 (85.7%)	63 (80.8%)	149 (82.8%)	83 (83.8%)
	Fair/ Poor/ Very poorly	24 (14.3%)	15 (19.2%)	31 (17.2%)	16 (16.2%)
RA		34 (20.1%)	9 (11.5%)	28 (15.3%)	15 (14.9%)
Frequency of medication use	Some/ Few/ No days	150 (91.5%)	68 (90.7%)	153 (84.1%)	86 (68.9%)
	All/ Most days	14 (8.5%)	7 (9.3%)	29 (15.9%)	13 (13.1%)

Footnote: n= number of observations; HADS= Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983); SD= Standard Deviation; ACR= American College of Rheumatology (Wolfe et al., 1990); BMI= Body Mass Index; RA= Rheumatoid Arthritis.

Participants whose hand state remained stable included a higher proportion who were in the youngest age group (65.4% of trajectory group), married or cohabiting (82.1%), employed (40.3%), and had received a hand operation (19.5%) compared to the other trajectories. However, a larger proportion of those who improved over the 6 years

(therefore moved into 'least affected' at 3 years or 6 years) were male (45.0%), did not have nodes (36.5%), had shorter symptom duration (46.6%) and did not frequently take medication for their hand symptoms (8.5%) compared to the other trajectories. They also had lower mean levels of depression symptoms (mean 3.96 (SD= 3.0)).

Deterioraters had a larger proportion of females (74.3%), those who lived alone (20.8%), had a routine/ manual occupation (49.7%), widespread pain (50.8%), sleep problems (54.1%), nodes (54.6%), longer symptom duration (67.2%), pain in two or more hands joints (54%) and took medication for their hand symptoms more frequently (15.9%) compared to the other three trajectories. They also had a higher mean level of depression symptoms (mean 4.65 (SD= 3.4)).

Participants in the fluctuaters trajectory had a higher frequency of those in the oldest age group (12.9%), were more likely to believe they had used their hands excessively in occupation or hobbies (83.0%) and had a lower proportion of bilateral hand pain (58.3%) compared to the other trajectories.

7.5.2 Intermediate baseline group - Multivariable model

Firstly, for the intermediate baseline group the same factors that were used as univariable factors in the four groups of *section 7.3* were again used to predict membership of the 6 year trajectories. The results of all the univariable analyses are presented in *Appendix G*. However, the multivariable model including all factors showing significant associations in the univariable analysis are displayed in *Table 7.5.2*.

Compared to membership in the 'improvers' trajectory, only individuals with nodes were significantly more likely to be 'stable' (RRR= 2.31 (1.24, 4.28)), although those with chronic hand pain duration were also more likely to be in the 'stable' trajectory, but this was not statistically significantly (RRR= 1.42 (0.76, 2.65), *p*-value= 0.27).

Table 7.5.2: Multivariable model of trajectories over 6 years follow-up: associations (RRR (95% CI)) for all significant univariable predictive factor.

Baseline factor (n= 449)		Improvers (n= 134)	Stable (n= 69)	Deterioraters (n= 156)	Fluctuaters (n= 90)
Gender	Male	1.00	1.05 (0.57,1.97)	**0.48 (0.28,0.83)**	0.86 (0.48,1.54)
ACR Widespread pain		1.00	1.20 (0.65,2.22)	1.24 (0.74,2.07)	1.27 (0.72,2.26)
HADS	Depression	1.00	1.02 (0.91,1.13)	1.02 (0.93,1.11)	**1.11 (1.01,1.22)**
BMI		1.00	1.02 (0.95,1.09)	0.96 (0.91,1.02)	1.06 (1.00,1.12)
Any sleep problems		1.00	0.96 (0.50,1.84)	**2.00 (1.18,3.39)**	0.79 (0.43,1.43)
Nodes		1.00	**2.31 (1.24,4.28)**	**1.86 (1.12,3.09)**	1.33 (0.74,2.38)
Previous 12 month hand pain duration	<3 months	1.00	1.00	1.00	1.00
	≥3 months	1.00	1.42 (0.76,2.65)	**1.79 (1.06,3.00)**	0.97 (0.55,1.72)
Bilateral hand pain		1.00	1.06 (0.55,2.02)	**1.85 (1.06,3.24)**	0.86 (0.48,1.57)
Pain in two or more hand joints	No/ Few days	1.00	1.00	1.00	1.00
	All/ Most/ Some days	1.00	1.06 (0.56,2.02)	1.13 (0.67,1.93)	1.01 (0.55,1.85)
Frequency of medication use	Some/ Few/ No days	1.00	1.00	1.00	1.00
	All/ Most days	1.00	0.72 (0.24,2.23)	1.75 (0.79,3.87)	1.68 (0.70,4.07)
Comorbidity		1.00	0.86 (0.47,1.58)	1.39 (0.84,2.32)	0.72 (0.41,1.28)

Footnote: RRR= Relative Risk Ratio; CI= Confidence Interval; n= number of observations; ACR= American College of Rheumatology (Wolfe et al., 1990); HADS= Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983); BMI= Body Mass Index; Bold entries represent significant at 5% after adjustment for other factors in table.

Factors significantly associated with membership in the 'deterioraters' trajectory were female gender (as male's RRR= 0.48 (0.28, 0.83)), sleep problems (RRR= 2.00 (1.18, 3.39)), nodes (RRR= 1.86 (1.12, 3.09)), chronic hand pain duration (RRR= 1.85 (1.06, 3.24)) and bilateral hand pain (RRR= 1.85 (1.06, 3.24)), compared to the 'improvers' trajectory. In addition, individuals who frequently took medication for their hand symptoms were more likely to be in the 'deterioraters' trajectory although this did not reach statistical significance (RRR= 1.75 (0.79, 3.87), *p*-value= 0.117).

Finally, individuals who were classified in the 'fluctuaters' trajectory were more likely to have a higher depression score (RRR= 1.11 (1.01, 1.22)) and have a higher BMI (RRR= 1.06 (1.00, 1.12), *p*-value= 0.058).

7.5.3 Worst baseline group - Baseline characteristics

There were three potential trajectories in which the participants who were in 'high pain and poor gross function' or 'severely affected' at baseline (worst groups) could be in, there are displayed in *Table 7.5.3*. 221 individuals (30% of analysis population) improved their hand state over the 6 year period, 103 fluctuated (14%), and 424 remained stable (57%).

Those in the 'improver' trajectory had a higher proportion in the oldest age group (14% of trajectory) compared to the other trajectories (8% for 'stable' and 13% for 'fluctuaters'). In addition, a larger proportion of individuals in the 'improvers' trajectory over the 6 years were male (35%), retired (55%) and had (previously) worked in a routine/ manual occupation (62%) compared to the proportions in other trajectories. A larger proportion of those in the stable trajectory were in the youngest age group (50-64 years, 57%) and had higher proportions of those who were female (74%), had (previously) worked in a higher managerial/ professional occupation (20%), had widespread pain (67%), sleep problems (57%), nodes (59%), chronic hand pain duration (85%), bilateral hand pain (82%), pain in two or more hand joints (84%), felt negatively about the impact of their hand problems compared to others of a similar age (48%), self-reported RA (21%) and frequently took medication for their hand symptoms (46%) compared to the other trajectories.

Table 7.5.3: Baseline characteristics of participants in each of the trajectories for those in the worst baseline groups (n (%) unless stated).

Baseline factor (n= 748)		Improvers (n= 221) (29.5%)	Stable (n= 424) (56.7%)	Fluctuaters (n= 103) (13.8%)
Age (years)	50-64	117 (52.9%)	240 (56.6%)	54 (52.4%)
	65-74	73 (33.0%)	149 (35.1%)	36 (35.0%)
	75+	31 (14.0%)	35 (8.3%)	13 (12.6%)
Gender	Male	78 (35.3%)	108 (25.5%)	35 (34.0%)
Lived alone		54 (24.9%)	101 (25.0%)	24 (25%)
Marital status	Married/ Cohabiting	151 (68.6%)	289 (68.7%)	68 (68.0%)
	Single	69 (31.4%)	132 (31.4%)	32 (32.0%)
Employment status	Retired	119 (54.8%)	205 (50.6%)	51 (52.6%)
	Employed	42 (19.4%)	73 (18.0%)	17 (17.5%)
	Others	56 (25.8%)	127 (31.4%)	29 (29.9%)
Social class	Higher managerial/ Professional	30 (14.6%)	80 (20.2%)	18 (19.8%)
	Intermediate	49 (23.8%)	99 (25.0%)	24 (26.4%)
	Routine/ Manual	127 (61.7%)	217 (54.8%)	49 (53.9%)
HADS Depression (mean (SD))		5.77 (3.7)	6.00 (3.8)	6.13 (4.1)
ACR Widespread pain		126 (57.0%)	284 (67.0%)	63 (61.2%)
Any sleep problems		106 (48.2%)	240 (56.7%)	54 (53.5%)
BMI (mean (SD))		27.7 (5.9)	28.1 (5.7)	27.6 (5.3)
Previous hand injury		68 (32.1%)	137 (33.3%)	27 (27.8%)
Hand operation		37 (17.2%)	89 (21.6%)	13 (13.4%)
Excessive use		186 (85.7%)	371 (88.3%)	93 (91.2%)
Nodes		81 (39.3%)	242 (59.3%)	49 (50.0%)
Previous 12 month hand pain duration	<3 months	55 (25.4%)	64 (15.2%)	21 (20.4%)
	≥3 months	162 (74.7%)	357 (84.8%)	82 (79.6%)
Bilateral hand pain		142 (62.1%)	345 (82.0%)	69 (67.0%)
Pain in two or more hand joints	No/ Few days	63 (29.7%)	68 (16.3%)	26 (25.7%)
	All/ Most/ Some days	149 (70.3%)	349 (83.7%)	75 (74.3%)
Impact of hand problem	Very well/ Well	135 (61.9%)	220 (52.5%)	61 (61.0%)
	Fair/ Poor/ Very poorly	83 (38.1%)	199 (47.5%)	39 (39.0%)
RA		34 (15.4%)	90 (21.2%)	17 (16.5%)
Frequency of medication use	Some/ Few/ No days	149 (70.3%)	227 (54.3%)	73 (72.3%)
	All/ Most days	63 (29.7%)	191 (45.7%)	28 (27.7%)

Footnote: n= number of observations; HADS= Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983); SD= Standard Deviation; ACR= American College of Rheumatology (Wolfe et al., 1990); BMI= Body Mass Index; RA= Rheumatoid Arthritis.

Participants classified in the 'fluctuaters' trajectory had slightly higher depression symptom scores (mean= 6.13 (SD= 4.1)) had a higher proportion with previous excessive hand use (91%) and a lower proportion who had had a hand injury (28%) or received a hand operation (13%) compared to the other two trajectories.

7.5.4 Worst baseline group - Multivariable model

The same process to determine factors associated with follow-up trajectories was performed on the participants who were in the 'worst' group at baseline ('high pain and poor gross function' or 'severely affected') as that completed in *section 7.5.2* for the intermediate groups. Again, the univariable results from the initial four groups are reported in *Appendix G*, and the multivariable model is displayed in *Table 7.5.4*. Similarly, the estimates are in reference to membership in the 'improvers' trajectory.

Table 7.5.4: Multivariable model of factors associated (RRR (95% CI)) with 6 year trajectory of individuals who were in the worst group at baseline.

Baseline factor (n= 661)		Improvers (n= 183)	Stable (n= 389)	Fluctuaters (n= 89)
Gender	Male	1.00	0.68 (0.45,1.03)	0.93 (0.52,1.64)
Age (years)	50-64	1.00	1.00	1.00
	65-74	1.00	0.82 (0.55,1.24)	0.90 (0.51,1.60)
	75+	1.00	**0.51 (0.27,0.94)**	1.02 (0.47,2.21)
ACR Widespread pain		1.00	1.21 (0.82,1.79)	1.29 (0.75,2.22)
Any sleep problems		1.00	1.13 (0.76,1.68)	0.96 (0.55,1.66)
SF-12 general health	Good/ Very good/ Excellent	1.00	1.00	1.00
	Poor/ Fair	1.00	1.33 (0.88,2.02)	0.97 (0.55,1.73)
Nodes		1.00	**1.83 (1.23,2.74)**	1.33 (0.76,2.33)
Previous 12 month hand pain duration	<3 months	1.00	1.00	1.00
	≥3 months	1.00	1.44 (0.89,2.33)	1.27 (0.65,2.45)
Bilateral hand pain		1.00	**1.64 (1.06,2.55)**	0.88 (0.49,1.57)
Pain in two or more hand joints	No/ Few days	1.00	1.00	1.00
	All/ Most/ Some days	1.00	1.31 (0.81,2.11)	1.27 (0.67,2.42)
Impact of hand problem	Very well/ Well	1.00	1.00	1.00
	Fair/ Poor/ Very poorly	1.00	1.26 (0.84,1.90)	1.06 (0.60,1.89)
Frequency of medication use	Some/ Few/ No days	1.00	1.00	1.00
	All/ Most days	1.00	1.42 (0.94,2.16)	0.90 (0.49,1.62)

Footnote: RRR= Relative Risk Ratio; CI= Confidence Interval; n= number of observations; ACR= American College of Rheumatology (Wolfe et al., 1990); SF-12= Short Form questionnaire (Ware et al. 1996); Bold entries represent significant at 5% after adjustment for other factors in table.

Participants in the ‘stable’ trajectory were more likely to be younger (75+ RRR= 0.51 (0.27, 0.94)) and more likely to have nodes (RRR= 1.83 (1.23, 2.74)) and bilateral hand pain (RRR= 1.64 (1.06, 2.55)). In addition to these factors, those who had chronic hand pain duration (RRR= 1.44 (0.89, 2.33)), negative self-perceived health status (RRR= 1.26 (0.84, 1.90)) and frequently took medication for their hand symptoms (RRR= 1.42 (0.94, 2.16)) were more likely to be in the ‘stable’ trajectory, but estimates did not reach statistical significance. No factors were found to be statistically associated with membership in the ‘fluctuaters’ trajectory (which due to the development of the trajectories in this more severe group, participants must have improved at 3 years, but then deteriorated at 6 years). However those with chronic hand pain duration (RRR= 1.27 (0.65, 2.45)), widespread pain (RRR= 1.29 (0.75, 2.22)) and nodes (RRR= 1.33 (0.76, 2.33)) were to some extent more likely to be classified as a ‘fluctuater’ than an ‘improver’.

7.6 Summary

This chapter explored whether baseline factors predicting long term outcomes, in addition to baseline hand state, could be identified. Age group, presence of sleep problems, nodes, chronic hand pain duration, bilateral hand pain and frequently taking medication were significantly associated with membership in more severe hand states at 3 years and 6 years. Generally similar factors (specifically age group, chronic hand pain duration, bilateral hand pain, nodes and sleep problems) were associated with membership in unfavourable 6 year trajectories (stable and deteriorated), as summarised in *Table 7.6.1*.

Table 7.6.1: Summary table of factors significantly associated (after multivariable adjustment) with hand state/ trajectory in each analysis.

Adjusted baseline factor associated with outcome	Associated with hand state at 3 and 6 years		Associated with trajectory over 6 years (improved, stable, fluctuated, deteriorated)	
	3 years	6 years	Intermediate groups	Worst groups
Gender		✓	✓	
Age	✓	✓		✓
Social class		✓		
ACR Widespread pain	✓			
HADS Depression	✓		✓	
Sleep problems	✓	✓	✓	
SF-12 general health	✓			
Nodes	✓	✓	✓	✓
Duration	✓	✓	✓	
Bilateral hand pain	✓	✓	✓	✓
Impact		✓		
Frequency of medication use	✓			
Baseline state	✓	✓	N.A.	N.A.

Footnote: ACR= American College of Rheumatology (Wolfe et al., 1990); HADS= Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983); BMI= Body Mass Index; SF-12= Short Form questionnaire (Ware et al., 1996).

7.6.1 Findings and relationship with previous literature

The factors that were consistently associated with a more unfavourable long-term outcome (presence of finger nodes, sleep problems, chronic hand pain duration, and bilateral hand pain) are discussed in the concluding chapter (*section 11.2.4*). However, factors that were unexpected (such as hand-specific factors that were not predictive of future hand state membership and the complex relationship between age and long-term hand state) are discussed here.

Whilst some hand-specific factors were strongly associated with poor long term outcomes (nodes, duration, bilateral pain, medication), other hand-specific factors, such as 'previous

hand injury', 'hand operation' and 'excessive use' (either in occupation or hobbies), were not independently associated with long term hand state (3 years, 6 years, or trajectory analyses) and RRR's were often close to 1. Due the nature of these factors, it might be expected that these factors would have a large impact on the pattern of hand pain/ problems (e.g. damaging to the hand sufficient for an operation or through overuse could be expected to lead to more rapid deterioration). However, this finding was similar to that found by Marshall and colleagues who assessed subsets of radiographic hand OA (such as nodal OA, erosive OA) in a small subset of the NorStOP population and investigated associations between 'previous hand operation' and 'excessive use' with membership of these subsets (Marshall et al., 2013). The authors found no significant association with these variables, similar to the findings of this study. The proportion of people reporting either of these factors was not particularly small (lowest n= 33 (15% for 'high pain' state for hand operation), all other states n>50 (>17%), *Appendix F*) for hand operation or previous injury, so it is unlikely this finding is explained by small frequency. A potential explanation for this could be due to recall bias of the participant, or more likely these factors could be predictive of the onset of hand pain/ problems, however they do not influence progression of the condition once individuals have it. This finding is supported in other previous studies, as these factors were not found to be associated with progression of hand pain/ problems (Spies-Dorgelo et al., 2008; Nicholls et al., 2012; Marshall et al., 2013), but have been found to be associated with onset (Caspi et al., 2001; Haara et al., 2003; Rossignol et al., 2005; Bernard et al., 2010).

There was some (but not consistent) evidence that older age was related to an improvement in outcomes. For example, those in the middle age group (65-74 years 'newly retired') were significantly less likely than those aged 50-64 years to belong to any of the problem function states at 3 or 6 years compared to being in the 'least affected' state, after

adjustment for baseline state. Although this finding was quite surprising, it could potentially be explained by the adjustment for baseline state. For example, without adjusting for baseline state, the older age groups were up to twice as likely to be in 'severely affected' state at 3 years (RRR= 1.58 (1.28, 1.95) for those aged 65-74 years, and RRR= 1.77 (1.31, 2.40) for those aged 75+ years, compared to those aged 50-64 years), and this was also the case in the 6 years analysis, although associations were slightly weaker. These results are partially explained by the fact that age group was strongly related to baseline state, in that younger participants were more likely to be in the less severe states at baseline, while the older participants were more likely to be in 'severely affected'. Therefore, at the follow-up points, many older participants can only improve or maintain stable states. A large amount of evidence has found a strong relationship between older age and the onset of hand problems (Caspi et al., 2001; Haara et al., 2003; Grotle et al., 2008a; Ghosh et al., 2014; Prieto-Alhambra et al., 2014), therefore age is potentially more strongly associated with the onset of hand pain/ problems than the progression of the condition.

There are three hypothesised explanations why age was not associated with the progression of hand pain/ problems, and why hand conditions appeared to improve for older individuals:

i) Hand conditions actually did improve over time;

ii) Hand conditions did not improve with age, but older people did not regard their condition as serious as it once was, especially in comparison to some of the other more life-threatening or disabling conditions that may have developed with age;

iii) Hand conditions were not necessarily improving, but due to individuals living with the condition for a number of years, they adapted their lifestyle to cope with the condition (which may include various gadgets to aid opening jars or the use of taps with levers for easy opening).

A PhD project on a subset of the NorStOP participants found that individuals with hand problems did report adapting their everyday lifestyle to cope with their condition, with 85% (n= 525) of the sample (n= 621) reported at least one adaption method (Myers, 2008).

Further potential evidence of this comes from two qualitative studies carried out on older people with OA. Firstly, Richardson and colleagues provided extracts of interviews describing how people with chronic joint pain develop resilience in dealing with their conditions, which could lead to an under-representation of individuals' problems over time, especially if the condition is seen as 'part of ageing' (Richardson et al., 2014). Secondly, a series of hand OA interviews revealed that many individuals find ways to adapt their lifestyle to cope with the condition (Bukhave and Huniche, 2014). Examples of this are electronic devices, such as toothbrushes, can openers and redesigning appliances around the home, such as handles, drawers and taps (Bukhave and Huniche, 2014).

Those who reported frequently taking medication at baseline were significantly more likely to be in unfavourable long-term hand states, particularly the 'high pain and poor gross function' state at 3 years (*Table 7.3.5*) and the 'high pain' state at 6 years (*Table 7.3.6*). Medication use is one of the main focuses of the next chapter, however briefly, in a qualitative study (Hill et al., 2011) participants with hand OA highlighted they were reluctant to take medication for their hand condition. For example, one participant in Hill's study stated they preferred to use homeopathic treatments (cod liver oil/ magnotherapy):

"it's far better than all these pills and you haven't got the side-effects to them" (page 1870, Hill et al., 2011).

Individuals who report taking medication may be those at a more severe stage of the condition and feel they require medication to ease their worsening symptoms. These

aspects are investigated further in *Chapter 8* where individuals who seek primary health care are explored.

7.6.2 Strengths and limitations

Knowledge of the factors that predict progression of hand pain/ problems in a primary care setting, and which factors are most associated with an unfavourable long-term outcome is limited (Nicholls et al., 2012). Therefore, a strength of these analyses is that this is one of the first studies to explore these issues. In addition to this, the long-term follow-up in this study (6 years) is longer than most previous progression studies, and contains a larger number of potential predictive factors (Spies-Dorgelo et al., 2008; Botha-Scheepers et al., 2009; Bijsterbosch et al., 2011).

The selection of predictive factors to include in a multivariable model based on univariable significance has been criticised, as relationships between variables that are not significant in a univariable analysis can yet be significant in a multivariable analysis (Sun et al., 1996). Therefore, a potential consequence in this study is that non-significant univariable factors were removed when they could have been significant in the multivariable analyses. However, this was an exploratory analysis as there was very little evidence for predictive factors of the long-term course of hand problems in older adults, and therefore this requires further research.

There are very few prospective cohort studies with a long-term follow-up (Nicholls et al., 2012) and most of these studies have focused on radiographic changes as an outcome (Marshall et al., 2013) or have included adults aged 18 years and older (Spies-Dorgelo et al., 2007; Spies-Dorgelo et al., 2008), while other studies have tended to focus on specific hand problems (e.g. carpal tunnel syndrome or rheumatoid arthritis (RA) (e.g. Uhlig et al., 2000; Nathan et al., 2002; Dyer et al., 2008; Giles et al., 2011; Burton et al., 2015)).

Therefore, this study used potential predictive factors that have been shown to be of importance across different musculoskeletal conditions, or have been found to be associated with hand pain/ problems in cross-sectional studies, including physical, psychological, and sociodemographic factors to start to investigate factors associated with poor outcomes in those with hand problems.

The trajectories defined and analysed in *sections 7.4* and *7.5* were defined manually and not using statistical methods (as the latent hand states were developed). While this approach does not use the data to generate the most representative trajectories of the study population, it was agreed with the supervision team that the trajectories developed made clinical sense and represented plausible pathways for potential patients to take. Due to the nature of using the two modified groups (rather than the four hand states) individuals who were switching between the 'high pain' and 'poor gross function' states over the 6 years were regarded as stable. Similarly in the 'worse' group, people switching between 'high pain and poor gross function' and 'severely affected' were considered to be stable. Transition probabilities between 'high pain' and 'poor gross function' were very low (all τ<0.031, according to the estimates produced in *Table 6.3.1*). However, transitions between 'high pain and poor gross function' and 'severely affected' were somewhat larger (τ ranges between 0.06 and 0.21), and therefore individuals switching between these two states would remain undetected due to the design of the trajectories. However, due to how similar these two states were in characteristics, this limitation was unlikely to result in different conclusions regarding the factors associated with trajectory membership.

The trajectory analysis also reduced the sample further from the main analysis, as analysis was completed in two groups (those in intermediate hand states and those in the worst hand states), therefore the number of observations in some trajectories were small (n= 78 for 'stable' in the intermediate group analysis for example). To ensure optimal accuracy of

estimates, the 'improvers' trajectory was chosen as the reference category in both trajectory analyses, because it was a reasonably large group (n= 169 and n= 221) compared to the observations in the other trajectories, it was available in both analyses and it makes logical sense to compare to the most favourable trajectory. A limitation of these analyses is the small sample size, which is reflected by the wide confidence intervals around the estimates. Consequently, discussion of results has avoided focussing solely on the significance of results and also explored the magnitude of associations.

The investigations contained in this chapter should be viewed as exploratory and hypothesis generating. Similarly, the factors found to be significant should be investigated in other datasets to validate their relationship to hand pain/ problems in older people.

7.6.3 Conclusions

Several factors have been revealed as potentially predictive of both the 3 years and 6 years hand state membership, with a number of these also predictive of 6 year hand trajectories. From the analyses presented in *Chapter 6* and *Chapter 7*, factors such as presence of finger nodes, sleep problems, chronic hand pain duration and bilateral hand pain appear to be associated with a more unfavourable outcome in older adults with hand problems. The implication of these findings will be discussed in *Chapter 11*; although they suggest that clinicians and researchers should consider these factors when treating patients and when investigating hand pain/ problems in future cohort studies.

The next chapter addresses the primary care records for hand conditions of the participants, in particular consultations and prescriptions, and assesses the association of the hand states with health care use.

Chapter 8: **Association of hand states with primary care consultations and medications**

8.1 Introduction and objectives

The focus of this chapter was to investigate the associations of the hand states developed in previous chapters with GP consultations and analgesia prescription. Previous work has found that those more likely to consult with hand (or wrist) problems were more likely to be female, have longer duration of symptoms, and recurrent problems; although this was carried out in consulters aged >18 years (Spies-Dorgelo et al., 2007).

Previous qualitative work within the Research Centre has identified that some patients with hand conditions report a lack of support from their doctor, and implied that as a result may not attend their GP for treatment (Hill et al., 2011). Analgesia prescription is part of the standard treatment approach for older people seeking health care for hand pain/ problems (NICE guidelines, 2014). However, there is little information regarding the outcome of primary care management for hand conditions.

Ideally, additional treatments aside from analgesia prescription would be assessed in this chapter, such as self-management advice and education, referrals to occupational therapists, physiotherapists, rheumatology or orthopaedics. However, only a small proportion of people with hand problems are referred to allied health professionals/ specialists (Dziedzic et al., 2007; Spies-Dorgelo et al., 2008) and referrals are often not well recorded in routinely collected health care data. As such it was decided to focus this chapter on analgesia only.

The specific objectives of the chapter were to:

i. Assess the extent of seeking primary health care and receiving prescription for hand problems;
ii. Determine associations of baseline hand state with consultation and prescription for hand problems;
iii. Descriptively compare 6 year hand states on receiving specific medications for hand problems.

8.2 Methods

8.2.1 NorStOP medical records review

Every participant who was mailed a HS questionnaire at baseline was asked whether or not they consented to medical record review. For participants who consented to medical record review, all available health care data was extracted from date of the baseline questionnaire through to the date of the 6 year survey. Naturally, this information contained information on all consultations/ prescriptions over the 6 year period, however only consultations related to the hand and relevant prescriptions were included here. Information in the medical records relevant to the objectives of this chapter included:

- Consultations: date of consultation, Read Code (described later in chapter), diagnostic term associated with each Read Code, bodily location of concern;
- Prescriptions: date of prescription, drug name, strength of drug, format of drug, BNF chapter (described later in methods).

8.2.2 Consultations

During a consultation in UK general practice reasons for attending are typically recorded using 'Read Codes'. Read Codes cover diagnoses and symptoms of conditions seen in general practice and are categorised in various chapters, where all codes within a chapter are related. For example, any Read code that starts 'N…' concerns a MSK diagnosis. More specifically, a Read Code of 'N05' is related to OA. This system can be particularly useful for research because it provides the ability to derive a collection of patients who are all consulting for similar conditions. Free text is also available to pair with each consultation, however, the content of this is determined by the GP and is therefore not uniformly completed.

Relevant consultations were identified from participants' medical records. These were consultations recorded with a Read Code related to 'hand' or 'hand/ wrist', and MSK in nature. Previous work within the Research Centre has developed approaches to identifying MSK consultations, including deriving MSK code lists for each bodily location (Jordan et al., 2010). Consistent with the definition of the study population, no restrictions were made regarding the specific recorded diagnosis or symptoms for consultations (such as restricting the consultations to those labelled as 'osteoarthritis' or 'hand pain'); however, consultations that were coded purely as 'wrist' problems were excluded from this investigation.

8.2.3 Prescriptions

Similar to consultations, any prescription written in the UK by a GP is categorised into a chapter, based on the British National Formulary (BNF, Joint Formulary Committee, 2015). The BNF is portioned into chapters, where each chapter refers to a particular system of the body, for example, 'chapter 1' relates to the gastro-intestinal system, 'chapter 2'

relates to the cardiovascular system, and so on. Within each chapter there are headings and sub-headings based on pharmacology and therapeutic use, and to further define drugs for a specific purpose. For example, ‘chapter 4’ is regarding the central nervous system, while ‘4.7’ relates to analgesia and ‘4.7.1’ relates to non-opioid analgesia.

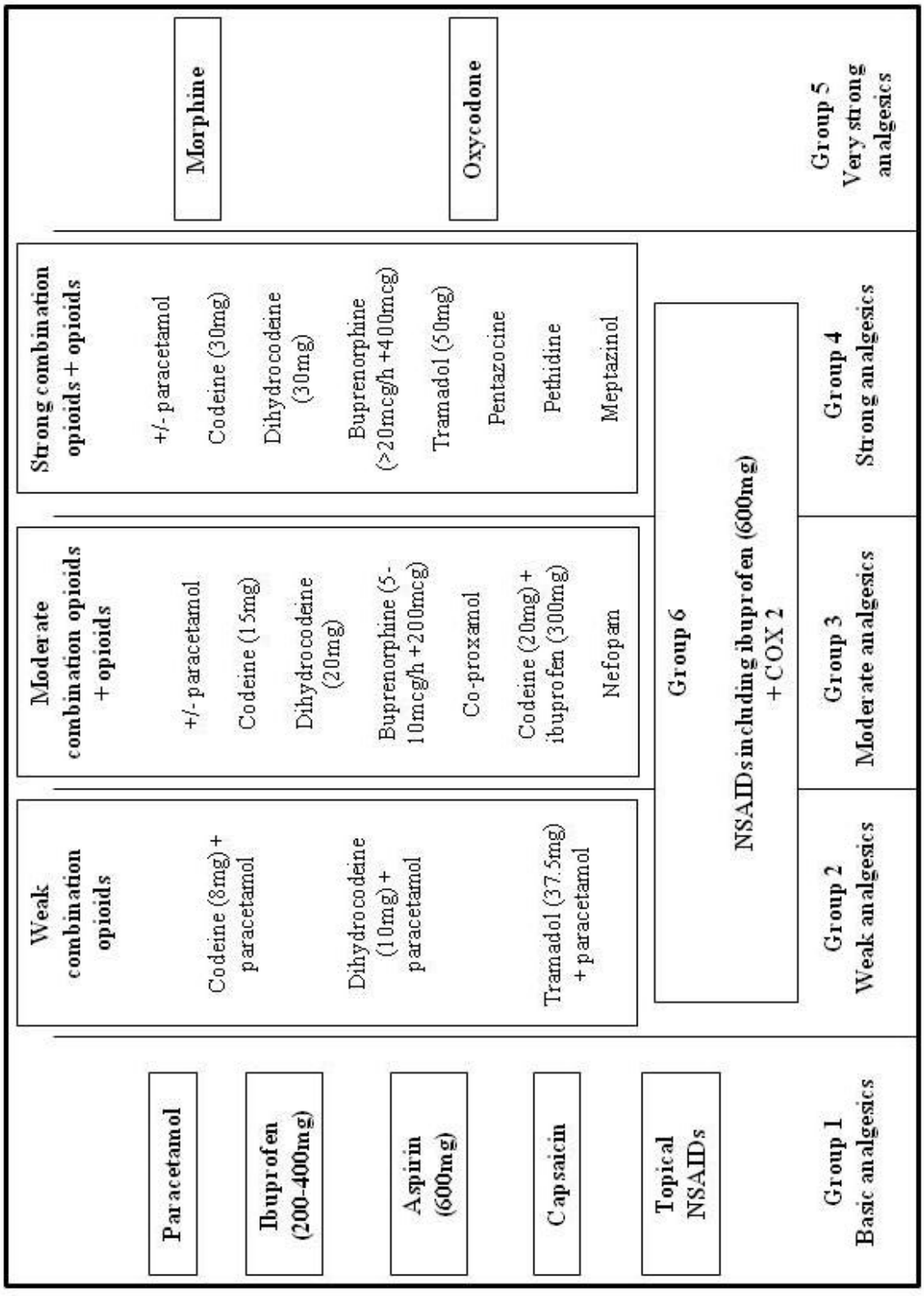
The drugs of interest for the prescription analysis are contained in sub-headings of ‘chapter 4’ (central nervous system), ‘chapter 10’ (MSK) and one drug in ‘chapter 5’ (anti-biotic). The sub-headings used are ‘4.7.1’ (non-opioid and compound analgesia), ‘4.7.2’ (opioid analgesia), ‘10.1.1’ (NSAIDs), ‘10.3.2’ (topical NSAIDs), and ‘5.1.8’ for Celecoxib (100mg/ 200mg) which although it is a NSAID, is classified in this sub-heading. These have been found to represent the potential analgesia drugs prescribed for joint-related problems (Bedson et al., 2013). In addition to this, all of the drugs prescribed within these sub-headings have also been classified into one of six groups; this is explained in the next sub-section (Bedson et al., 2013).

The reason for a prescription is rarely directly recorded, so it was agreed with supervisors that for an analgesia prescription to be considered related to hand problems, the participant had to have consulted for a hand-related problem within 14 days prior to the date of prescription (so on the same day, or within the previous 2 weeks). This approach is consistent with other previous work at the Research Centre (Edwards et al., 2015).

8.2.4 Classification of analgesia

All potential prescriptions for MSK joint pain have previously been identified and classified into six groups that represent various strengths/ potencies (Bedson et al., 2013). This developed model is displayed in *Figure 8.2.1*, where some examples of common drugs found in each group are presented.

Figure 8.2.1: Categorisation of potential MSK analgesia and NSAID prescriptions (adapted with permission from Bedson et al., 2013).



Footnote: N.B. 'Co-proxamol' has been withdrawn from prescribing in 2005 but was available in the first few years covered by prescription analysis.

Groups one to five represent increasing strengths of analgesia, from simple paracetamol and compound analgesia through to morphine and oxycodone. Group six represents NSAIDs which have an anti-inflammatory and analgesia effect and so do not fit into the linear potency hierarchy, and therefore group six is presented spanning various groups (for example, participants with MSK conditions may receive prescriptions from both group six and group two).

In addition to assessing whether participants received any analgesia prescription or not, three further analyses were performed to investigate associations with specific prescriptions: NSAIDs, mild analgesia and moderate/ strong analgesia. The first of these assessed whether participants were prescribed a drug from group six, which represented NSAIDs. The second concerned those who were prescribed a drug from group one or group two, which represent mild analgesia (and medication that can be purchased over the counter). The third group included participants who received a prescription from groups three, four or five, which contains moderate and strong analgesia.

8.2.5 Analysis

The analysis population for each comparison was hierarchical, as consultations for hand problems were assessed for all participants who consented to medical record review, and the prescriptions analysis was performed on only participants who consulted. The population used in each analysis is displayed in the results section (*Figure 8.3.1*).

Firstly, a descriptive assessment of the differences between the participants who did and did not consent to medical record review was performed in order to explore whether there were any substantial differences between the two groups. This assessment included the demographic and general health variables: age, gender, marital status, living status, employment status, social class, anxiety and depression score, widespread pain, sleep

problems, BMI, other comorbidities and baseline hand state (through frequency/ percentage, or mean/ standard deviation as appropriate). Secondly, for those who consented to medical record review, the participants who consulted for hand-related problems at any point over the 6 years (baseline to 6 year survey date) were compared descriptively to those who did not in terms of hand state at baseline and 6 years.

Subsequent to this, logistic regression was used in three separate analyses for both consultation and prescription associations, and to explore whether hand state predicted consultation or prescription status independently of demographic factors, and explore which hand-specific factors potentially explain the association of hand state with consultation/ analgesia prescription. Firstly, univariable logistic regression was used to investigate the association between baseline hand state and consulting for a hand-related condition over the 6 year period. Secondly, a multivariable analysis (model A) was computed to explore the relationship between baseline hand state and consultation after adjustment for demographic factors (age, gender, social class). Finally, a further multivariable analysis (model B) was carried out with the additional adjustment of the hand factors included in group three of *Chapter 7* (*section 7.3.3*). However, three factors (previous hand injury, previous hand operation, previous diagnosis of RA) which could plausibly explain individuals consulting their GP, were not included in the adjustment. Therefore, the hand factors included in model B were: excessive use (in occupation or hobby), nodes, duration, bilateral hand pain, impact of hand problems and frequency of medication use for hand symptoms.

The same process was also used to assess participants who received a prescription within 2 weeks of a hand or hand/ wrist consultation, within the population of those who consulted; the same factors were assessed, including baseline hand state, whist also exploring the impact of removing the factor of frequent medication use for hand symptoms in model B.

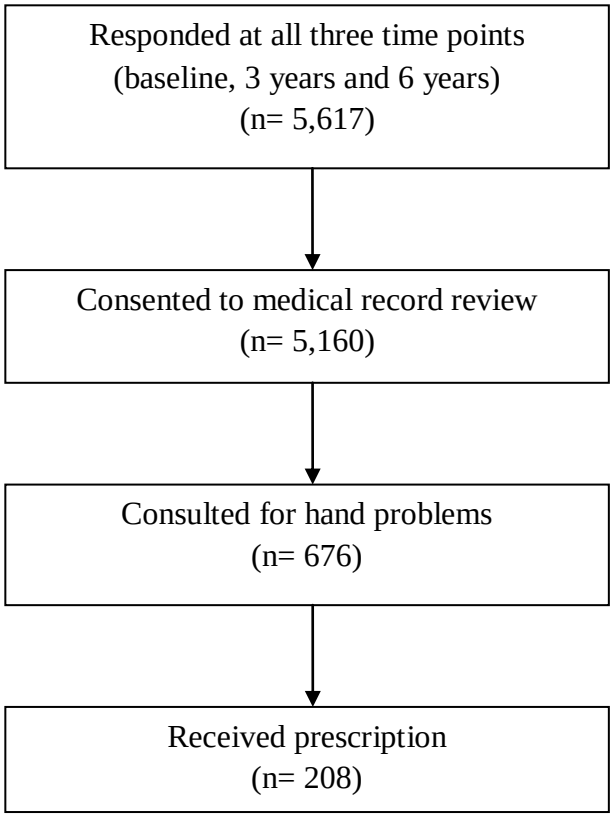
As a final analysis, hand state at baseline and 6 years was assessed for participants who received specific prescriptions (NSAIDs, mild analgesia, and moderate/ strong analgesia) in participants who received a prescription. Due to the small number of individuals receiving these specific prescriptions, these investigations were assessed descriptively by determining frequency and percentage of individuals receiving each prescription type by hand state.

For each of the comparisons statistical significance was assessed for categorical variables using a chi-square (χ^2) test, or a t-test for continuous variables to test for differences between consultation/ prescription status (yes/ no) in each analysis, and with logistic regression for univariable and adjusted analyses. Statistical significance was regarded as *p*-value<0.05. STATA version 13.1 was used for the analysis in this chapter (StataCorp, 2013).

8.3 Results

The analysis population for each stage of analysis is displayed in *Figure 8.3.1*. Of the 5,617 participants that responded at baseline, 5,160 (91.9%) consented to medical record review. Of this population, 676 (13.1%) consulted for hand or hand/ wrist problems over the 6 year period. Of the consulters, 208 (30.8%) received an analgesia prescription within 14 days of a hand-related consultation over the 6 year period.

Figure 8.3.1: Analysis population for each stage of the medical records analysis.



8.3.1 Consenters

There were a few differences between participants who did and did not consent to medical record review (*Table 8.3.1*). A higher proportion of males consented to review (94%) compared to females (91%), whereas higher proportions of older people consented to review (75+ years: 94% vs. 50-64 years: 91%). A larger proportion of individuals consenting to review were married compared to single (92% vs. 90%), retired compared to other (93% vs. 90%) and consenters had lower HADS anxiety scores (mean 6.2 vs. 6.7). Participants in the 'severely affected' state at baseline were more likely to consent.

Table 8.3.1: Description of hand states and baseline characteristics for participants who did and did not consent to medical record review (n (%), unless stated otherwise).

Responded at all time points (n= 5,617)		Did not consent	Did consent
Observations (n)		457	5,160
Baseline state	Least affected	356 (8.2%)	3,982 (91.8%)
	High pain	14 (6.3%)	210 (93.8%)
	Poor gross function	25 (8.1%)	282 (91.9%)
	High pain & PGF	43 (10.9%)	351 (89.1%)
	Severely affected	19 (5.4%)	335 (94.6%)
3 year state	Least affected	326 (8.0%)	3,742 (92.0%)
	High pain	26 (8.3%)	287 (91.7%)
	Poor gross function	23 (9.4%)	223 (90.7%)
	High pain & PGF	49 (9.2%)	486 (90.8%)
	Severely affected	33 (7.3%)	422 (92.8%)
6 year state	Least affected	343 (8.7%)	3,623 (91.4%)
	High pain	20 (6.4%)	292 (93.6%)
	Poor gross function	22 (9.5%)	209 (90.5%)
	High pain & PGF	38 (6.5%)	548 (93.5%)
	Severely affected	34 (6.5%)	488 (93.5%)
Gender	Female	288 (9.5%)	2,743 (90.5%)
	Male	169 (6.5%)	2,417 (93.5%)
Age (years)	50-64	325 (9.5%)	3,093 (90.5%)
	65-74	102 (6.2%)	1,557 (93.9%)
	75+	30 (5.6%)	510 (94.4%)
Marital state	Married/ Cohabiting	325 (7.7%)	3,914 (92.3%)
	Single	130 (9.7%)	1,205 (90.3%)
Lived alone	Did not live alone	342 (7.8%)	4,046 (92.2%)
	Lived alone	97 (9.5%)	923 (90.5%)
Employment status	Retired	181 (7.1%)	2,368 (92.9%)
	Employed	183 (9.0%)	1,859 (91.0%)
	Other	86 (9.7%)	805 (90.4%)
Social class	Higher managerial/ Professional	95 (7.0%)	1,267 (93.0%)
	Intermediate	106 (7.5%)	1,313 (92.5%)
	Routine/ Manual	224 (8.8%)	2,330 (91.2%)
HADS (mean (SD))	Anxiety	6.68 (4.0)	6.23 (4.0)
	Depression	3.92 (3.2)	3.89 (3.2)
ACR Widespread pain	No	335 (8.0%)	3,857 (92.0%)
	Yes	122 (8.6%)	1,303 (91.4%)
Any sleep problems	No	318 (8.6%)	3,374 (91.4%)
	Yes	138 (7.3%)	1,752 (92.7%)
BMI (mean (SD))		26.53 (4.3)	26.82 (4.5)
Comorbidity	No	242 (8.4%)	2,655 (91.7%)
	Yes	215 (7.9%)	2,505 (92.1%)

Footnote: n= number of observations; PGF= 'poor gross function'; HADS= Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983); SD= Standard Deviation; ACR= American College of Rheumatology (Wolfe et al., 1990); BMI= Body Mass Index.

8.3.2 Consulters

676 participants (13.1% of consenters) consulted their GP over the 6 year period for a hand or hand/ wrist problem. In the initial unadjusted analysis, all participants that consulted over the 6 year period were significantly more likely to be in a problem state (i.e. not in 'least affected') at either time point (baseline or 6 years) (*Table 8.3.2*). At 6 years, there were twice as many participants that had consulted in the 'high pain and poor gross function' and 'severely affected' states than those in the 'least affected' hand state. Of the 335 individuals who were classified in the 'severely affected' hand state at baseline, 74 (22%) consulted for hand problems over the next 6 years (*Table 8.3.2*). Similarly, of the 488 individuals classified in the 'severely affected' hand state at 6 years, 112 (23%) had consulted in the prior 6 years.

Table 8.3.2: Comparison of hand states at baseline and 6 years between participants who did and did not consult for hand-related problems in participants who consented to medical record review.

Consenters (n= 5,160)		Did not consult	Did consult	*p*-value
Observations (n)		4,484	676	
Baseline state	Least affected	3,527 (88.6%)	455 (11.4%)	$p<0.001$
	High pain	173 (82.4%)	37 (17.6%)	
	Poor gross function	237 (84.0%)	45 (16.0%)	
	High pain & PGF	286 (81.5%)	65 (18.5%)	
	Severely affected	261 (77.9%)	74 (22.1%)	
6 year state	Least affected	3,266 (90.2%)	357 (9.9%)	$p<0.001$
	High pain	241 (82.5%)	51 (17.5%)	
	Poor gross function	176 (84.2%)	33 (15.8%)	
	High pain & PGF	425 (77.6%)	123 (22.5%)	
	Severely affected	376 (77.1%)	112 (23.0%)	

Footnote: n= number of observations; PGF= 'poor gross function'.

Univariable analysis indicated the more severe hand states were more strongly associated with consultation over the 6 years follow-up ('high pain and poor gross function' Odds Ratio (OR)= 1.8 (1.3, 2.4), 'severely affected' OR= 2.2 (1.7, 2.9)) compared to those in 'least affected' at baseline (*Table 8.3.3*).

Table 8.3.3: Univariable and multivariable results (OR (95% CI)) of factors associated with consulting for hand problems over a 6 year period in participants who consented to medical record review.

Baseline factor (n= 5,160)		Univariable	Multivariable A	Multivariable B
Baseline hand state	Least affected	1.00	1.00	1.00
	High pain	1.66 (1.15, 2.40)	1.65 (1.13, 2.41)	1.19 (0.78, 1.84)
	Poor gross function	1.47 (1.05, 2.05)	1.41 (1.00, 2.00)	0.84 (0.55, 1.28)
	High pain & PGF	1.76 (1.32, 2.35)	1.72 (1.28, 2.31)	1.00 (0.68, 1.46)
	Severely affected	2.20 (1.67, 2.90)	2.15 (1.61, 2.89)	1.17 (0.75, 1.83)
Gender	Male	0.72 (0.61, 0.85)	0.78 (0.66, 0.93)	0.64 (0.49, 0.85)
Age (years)	50-64	1.00	1.00	1.00
	65-74	0.92 (0.76, 1.10)	0.89 (0.73, 1.07)	0.85 (0.64, 1.11)
	75+	0.90 (0.68, 1.20)	0.95 (0.70, 1.27)	0.58 (0.35, 0.97)
Social class	Higher managerial/ Professional	1.00	1.00	1.00
	Intermediate	1.03 (0.81, 1.30)	1.00 (0.78, 1.26)	1.09 (0.76, 1.57)
	Routine/ Manual	1.18 (0.96, 1.46)	1.08 (0.88, 1.34)	1.29 (0.94, 1.78)
Previous hand injury*		1.23 (0.97, 1.55)		
Hand operation*		1.69 (1.29, 2.21)		
Excessive use		0.84 (0.65, 1.08)		0.70 (0.52,0.96)
Nodes		1.54 (1.24, 1.90)		1.30 (1.01, 1.69)
Previous 12 month hand duration	≥3 months	1.29 (1.02, 1.62)		1.15 (0.87, 1.52)
Bilateral hand pain		1.13 (0.90, 1.43)		0.98 (0.75, 1.28)
Impact of hand problem	Fair/ Poor/ Very poorly	1.23 (0.95, 1.60)		1.17 (0.84, 1.63)
RA*		1.15 (0.86, 1.55)		
Frequency of medication use	All/ Most days	1.51 (1.14, 1.99)		1.21 (0.84, 1.73)

Footnote: OR= Odds Ratio; CI= Confidence Interval; n= number of observations; model A: adjusted for demographic factors (age, gender, social class); model B: additionally adjusted for excessive use, nodes, duration, bilateral hand pain, impact and frequency of medication use; PGF= 'poor gross function'; * not included as confounder in any analysis; RA= Rheumatoid Arthritis.

After adjustment for age, gender and social class, the association between baseline hand state and consultation remained. However after adjusting for the hand factors listed in model B, the associations were significantly weakened and close to 1 (null effect). Therefore, while a relationship was seen between hand state and consultation for hand problems in the univariable analysis and also after adjusting for demographic factors, the relationship was not independent of hand-specific factors (adjusting for these factors attenuated the association). The hand-specific factors of excessive use, frequent medication use for hand symptoms, and particularly presence of nodes, appeared strongly related to consulting for hand problems (*Table 8.3.3*).

8.3.3 Prescriptions

208 participants (30.8% of the consulters) received an analgesia prescription within 14 days of a hand problem consultation over the 6 year period. Similar to the analysis of consultations, a higher proportion of participants in problem hand states received a prescription during follow-up, with the exception of individuals in the 'poor gross function' state at baseline (who reported few pain items), compared to consulters in the 'least affected' hand state (*Table 8.3.4*).

Table 8.3.4: Comparison of hand states between participants who did and did not receive a prescription for hand problems, in participants who consulted for a hand-related problem.

Consulters (n= 676)		Did not receive a prescription	Did receive a prescription	*p*-value
Observations		468	208	
Baseline state	Least affected	325 (71.4%)	130 (28.6%)	*p*= 0.013
	High pain	24 (64.9%)	13 (35.1%)	
	Poor gross function	37 (82.2%)	8 (17.8%)	
	High pain & PGF	41 (63.1%)	24 (36.9%)	
	Severely affected	41 (55.4%)	33 (44.6%)	
6 year state	Least affected	273 (76.5%)	84 (23.5%)	*p*= 0.001
	High pain	29 (56.9%)	22 (43.1%)	
	Poor gross function	20 (60.6%)	13 (39.4%)	
	High pain & PGF	78 (63.4%)	45 (36.6%)	
	Severely affected	68 (60.7%)	44 (39.3%)	

Footnote: n= number of observations; PGF= 'poor gross function'.

Again, similar to the consultation analysis, participants who received a prescription were more likely to be in 'severely affected' than 'least affected' state (45% vs. 29% at baseline, 39% vs. 24% at 6 years). Hand state at baseline appeared to be associated with receipt of a prescription for participants with hand pain complaints (ORs> 1.3), but significantly only for individuals in the 'severely affected' state (OR= 2.0 (1.2, 3.3)) (*Table 8.3.5*). The strength of this association was similar when adjusting for demographic factors (model A). Adjustment for the hand factors (model B) reduced the OR substantially (from 2.0 (1.2, 3.4) to 1.3 (0.6, 3.2)), however an association was still present.

Individuals with bilateral hand pain were significantly less likely to receive a prescription in the multivariable model (OR= 0.6 (0.3, 1.0)). Participants with ≥3 months hand pain duration were significantly more likely to receive a prescription over the 6 year period (OR= 2.1 (1.2, 3.9)), as were those who reported frequently taking medication for their hand symptoms (OR= 2.3 (1.3, 4.6)). Adjusting for the factor 'frequent medication use' could be regarded as an over-adjustment (considering its direct link to the outcome of receiving a prescription). Removing this variable from the analysis in model B changed the estimate for participants 'severely affected' to 1.79 (0.8, 4.0, *p*-value= 0.16); no other major changes were observed for any of the other estimates.

Table 8.3.5: Univariable and multivariable results (OR (95% CI)) of factors associated with receiving a prescription for hand problems, in participants who consulted over a 6 year period for a hand-related problem.

Baseline factor (n= 676)		Univariable	Multivariable A	Multivariable B
Baseline hand state	Least affected	1.00	1.00	1.00
	High pain	1.35 (0.72, 2.74)	1.43 (0.68, 2.99)	1.02 (0.42, 2.47)
	Poor gross function	0.54 (0.25, 1.19)	0.53 (0.23, 1.18)	0.46 (0.18, 1.22)
	High pain & PGF	1.46 (0.85, 2.52)	1.39 (0.78, 2.46)	1.32 (0.62, 2.81)
	Severely affected	2.01 (1.22, 3.32)	1.98 (1.15, 3.41)	1.33 (0.56, 3.17)
Gender	Male	0.72 (0.51, 1.01)	0.80 (0.55, 1.18)	0.83 (0.45, 1.53)
Age (years)	50-64	1.00	1.00	1.00
	65-74	1.63 (1.14, 2.33)	1.83 (1.26, 2.68)	1.29 (0.75, 2.25)
	75+	0.96 (0.53, 1.75)	1.23 (0.66, 2.28)	1.20 (0.42, 3.46)
Social class	Higher managerial/ Professional	1.00	1.00	1.00
	Intermediate	1.39 (0.86, 2.24)	1.23 (0.75, 2.03)	1.29 (0.60, 2.79)
	Routine/ Manual	1.07 (0.70, 1.63)	0.93 (0.59, 1.46)	0.89 (0.44, 1.79)
Previous hand injury*		0.91 (0.58, 1.42)		
Hand operation*		0.69 (0.40, 1.18)		
Excessive use		1.39 (0.83, 2.33)		1.76 (0.93, 3.34)
Nodes		1.21 (0.79, 1.85)		1.25 (0.73, 2.11)
Previous 12 month hand duration	≥3 months	2.49 (1.53, 4.05)		2.11 (1.16, 3.85)
Bilateral hand pain		0.73 (0.47, 1.14)		0.56 (0.32, 0.97)
Impact of hand problem	Fair/ Poor/ Very poorly	1.17 (0.71, 1.92)		0.52 (0.26, 1.06)
RA*		1.07 (0.60, 1.93)		
Frequency of medication use	All/ Most days	2.19 (1.31, 3.64)		2.28 (1.13, 4.62)

Footnote: OR= Odds Ratio; CI= Confidence Interval; n= number of observations; model A: adjusted for demographic factors (age, gender, social class); model B: additionally adjusted for excessive use, nodes, duration, bilateral hand pain, impact and frequency of medication use; PGF= 'poor gross function'; * not included as confounder in any analysis; RA= Rheumatoid Arthritis.

8.3.4 Type of analgesia

Table 8.3.6 displays the frequency and proportion of individuals who received analgesia by type of analgesia in participants who consulted, stratified by baseline hand state. Approximately a fifth of individuals (n= 142, 21%) who consulted were prescribed a drug classified in the 'mild analgesia' group, such as paracetamol and co-codamol, with similar proportions in each baseline hand state, although lower in the 'poor gross function' hand state. Around one in 10 consulters received a NSAID over the 6 year period (n= 70, 10%), while only 24 participants (3.6%) were prescribed strong analgesia.

Table 8.3.6: Frequency and proportion of individuals who received specific analgesia in participants who consulted for a hand-related problem.

Consulters (n= 676)		Prescribed		
		NSAID	Mild analgesia	Moderate/ Strong analgesia
Baseline state	Least affected (n= 455)	45 (9.9%)	93 (20.4%)	11 (2.4%)
	High pain (n= 37)	4 (10.8%)	9 (24.3%)	3 (8.1%)
	Poor gross function (n= 45)	5 (11.1%)	5 (11.1%)	0 (0.0%)
	High pain & poor gross function (n= 65)	9 (13.9%)	17 (26.2%)	1 (1.5%)
	Severely affected (n= 74)	7 (9.5%)	18 (24.3%)	9 (12.2%)

Footnote: n= number of observations; NSAID= Non-steroidal anti-inflammatory drug.

The next analysis used the population that received a prescription (n= 208), and as explained in the methods section (*8.2.5*), this was only a descriptive analysis. There was little overlap between prescriptions for each of the specific analgesia, as the largest group received a combination of mild analgesia and NSAIDs over the 6 year period (n= 16, 7.7% of those who consulted and received a prescription), and only one individual was prescribed a drug from all three specific analgesia groups.

70 participants (33.7%) who consulted and received any prescription (n= 208), received a prescription for NSAIDs (*Table 8.3.7*). Amongst the individuals in the 'severely affected' hand state at baseline (and consulted and received a prescription), 21% received an NSAID

prescription, 55% received a mild analgesia prescription, and 27% received a moderate/ strong analgesia prescription over the subsequent 6 years (*Table 8.3.7*), whilst one person received both NSAIDs and mild analgesia. The proportion of individuals receiving NSAIDs and mild analgesia was higher within all of the less severe hand states at baseline (not for moderate/ strong analgesia).

Of the individuals in the 'severely affected' hand state at 6 years 32% received an NSAID prescription, 61% received a mild analgesia prescription, and 21% received a moderate/ strong analgesia prescription in the prior 6 years. In general, the proportion of individuals receiving mild analgesia increased between baseline and 6 years for all hand states, while NSAID and moderate/ strong analgesia prescriptions were comparable (in those who consulted and received a prescription) (*Table 8.3.7*).

Table 8.3.7: Comparison of hand states at baseline and 6 years between participants who did and did not receive specific analgesia, in participants who consulted and received a prescription.

		Prescribed		
Consulted & prescribed (n= 208)		NSAID	Mild analgesia	Moderate/ Strong analgesia
Observations		70	142	24
Baseline state	Least affected	45 (34.6%)	93 (71.5%)	11 (8.5%)
	High pain	4 (30.8%)	9 (69.2%)	3 (23.1%)
	Poor gross function	5 (62.5%)	5 (62.5%)	0 (0%)
	High pain & PGF	9 (37.5%)	17 (70.8%)	1 (4.2%)
	Severely affected	7 (21.2%)	18 (54.6%)	9 (27.3%)
6 year state	Least affected	29 (34.5%)	57 (67.9%)	6 (7.1%)
	High pain	6 (27.3%)	15 (68.2%)	2 (9.1%)
	Poor gross function	4 (30.8%)	9 (69.2%)	1 (7.7%)
	High pain & PGF	17 (37.8%)	34 (75.6%)	6 (13.3%)
	Severely affected	14 (31.8%)	27 (61.4%)	9 (20.5%)

Footnote: n= number of observations; NSAID= Non-steroidal anti-inflammatory drug; PGF= 'poor gross function'.

8.4 Summary

This chapter explored the association of the hand states (as identified in *Chapters 4* to *6*) with GP consultation and analgesia prescription for hand pain/ problem. Only a small proportion (676/ 4,484 (13.1%)) of people consulted their GP over a 6 year period with

hand problems, with those reporting more severe problems most likely to consult. However, only 22% of individuals classified in the 'severely affected' hand state at baseline consulted over the 6 year period. A third of participants who consulted received any analgesia prescription, with the majority of participants receiving mild analgesia (n= 142, 65%), some receiving NSAIDs (n= 70, 32%), and very few receiving moderate to strong analgesia (n= 24, 11%). Participants in the 'severely affected' state were most likely to be prescribed moderate to strong analgesia.

As expected, a larger proportion of participants consulting with hand problems were in the more severe hand states at either time point compared to those who did not consult. However, a tenth (10-11%) of participants who did consult for hand problems were actually in the 'least affected' states at either time point (357 participants at baseline, 455 at 6 years). It may be that a proportion of these individuals experienced episodes of hand pain/ problems during the 6 years but not at the time of the surveys.

Baseline hand state was predictive of consultation for hand pain/ problems over the 6 year period, even after adjustment for demographic factors. However, after adjustment for hand-specific factors, the association attenuated. Similarly, individuals in the 'severely affected' hand state at baseline were still significantly more likely to receive an analgesia prescription over the 6 year period after adjustment for demographic factors, however adjusting for hand-specific factors removed the association. A potential reason for this is that the hand-specific factors are likely to reflect reasons for consultation and prescription of analgesia for people with hand pain/ problems, therefore minimising any association.

A further interesting finding was in participants who did consult over the 6 year period, analgesia prescription did not appear to be determined by hand state severity at either baseline or 6 years. The association between receiving (any) analgesia and hand state at

both time points was significant in participants who consulted (p= 0.013 at baseline, p= 0.001 at 6 years, *Table 8.3.4*); however, there was not a clear trend with increased analgesia prescription and hand state severity (approximately a third of those in each of 'least affected', 'high pain' and 'high pain and poor gross function' at baseline received a prescription over the 6 years). Additional research is required to determine whether GP prescribing for hand conditions are based on patient pain and function symptoms, GP preference or other factors (such as comorbid conditions and polypharmacy (Roberts et al., 2014)).

Previous qualitative work has revealed that some patients with hand problems lose faith that their GP can treat their hand problem and have expressed opinions that doctors cannot help (Hill et al., 2011). For example, some patients were reported to say:

"I went to the GP. He said, 'no, there' nothing we can do with that" (page 1868, Hill et al., 2011);

"I think another frustrating thing is that the doctors just 'oh it's arthritis' and that's it" (page 1868, Hill et al., 2011);

These comments highlight that patients feel doctors are not able to do anything to help (detailed in Hill et al., 2011). An alternative explanation to this could be that older individuals see their hand condition as part of 'ageing' (discussed in *section 7.6.1*), and therefore do not regard seeking medical help as necessary. The work contained in this chapter found a low consultation rate over the 6 year period (13%) which does imply a reluctance of individuals to consult their GP, especially as 23% reported hand pain and/ or function symptoms at baseline alone.

Limitations relating to the analyses presented in this chapter are discussed in the main discussion chapter (*section 11.3.6*), along with the general limitations of the project. The next chapter explores the statistical technique of ALTA. This technique investigated whether changes in hand state over the 6 year period resulted in changes in mental health; this was achieved by exploring the longitudinal association between the latent hand states previously developed and used throughout this thesis, and a new set of latent states based on mental health (from items in the HADS).

Chapter 9: **Longitudinal associations between hand state and mental health**

9.1 Introduction and objectives

Direct associations between mental health and MSK conditions have consistently been shown in previous literature (for example, Bair et al., 2003; Arola et al., 2010; Kroenke et al., 2011). Research has shown that bodily pain was a significant predictor of developing anxiety and depression over a 3 year follow-up period in a subset of the NorStOP population (Arola et al., 2010). In addition to this, baseline anxiety and depression were also significant predictors of developing bodily pain (Arola et al., 2010). Further studies have shown associations of depression and anxiety with MSK conditions at specific sites including the knee (Mallen et al., 2007; Thomas et al., 2008), neck (Demyttenaere et al., 2007) and back (Demyttenaere et al., 2007).

However, there is limited research on whether hand conditions follow the general observed patterns of longitudinal associations of mental health with MSK problems, despite qualitative and cross-sectional evidence regarding the high impact of hand pain on ability to manage everyday activities (Hill et al., 2007; Spies-Dorgelo et al., 2007; Hill et al., 2011). Previous findings in this thesis have found that individuals in more severe hand states had more contemporaneous symptoms of anxiety and depression (*section 6.2*), while some mild association was found in the exploratory analyses of long-term hand state (*sections 7.3.5* and *7.5.2*). In addition to this, individuals in the RUG group meeting (discussed in *Chapter 4*, *sections 4.4.3* and *4.5.2*) also reported that the impact of hand

problems in everyday life may affect mental health. For example, one of the participants reported struggling to pick up their grandchildren due to a lack of hand function, and other activities that could potentially affect mental state. If the presence of hand pain/ problems is associated with poorer outcomes of mental health, then it is of clinical importance and should be investigated further. It may be possible to identify longitudinal associations between mental health and hand pain/ problems, through a novel extension to LTA, Associative Latent Transition Analysis (ALTA), which can model two potentially related areas or domains over a period of follow-up (Flaherty, 2008). A search of the literature (*Chapter 3*) identified only three publications using ALTA. These explored relationships between psychological state with substance use, psychological state with alcohol dependency, and alcohol use and sexual risk (Flaherty, 2008; Witkiewitz and Villarroel, 2009; Bray et al., 2010).

The specific objectives of this chapter were to:

i. Develop longitudinal phenotypes of mental health over the 6 years follow-up;
ii. Explore whether mental health state was associated with concurrent and previous hand state, using the novel longitudinal approach of ALTA.

9.2 Mental health phenotype development using LTA

The first part of the ALTA process is to develop the phenotypes for each domain separately using LTA. The base model developed in *Chapter 4* to *Chapter 6* was used for the hand states. The overall purpose of this sub-section, therefore, was to develop mental health phenotypes from items in the HADS using the same study population that was used to create the hand phenotypes in *Chapter 4*, and using the same indicator selection process. The patterns of symptoms regarding anxiety and depression in this population is described, and rates of transitions between the various states of mental health over the 6 year period

identified. The same methodology used for developing the hand phenotypes was applied (e.g. for item selection and restricting item-response probabilities over time); this is therefore only briefly presented, but the full results from the model selection process are displayed in *Appendix H*.

9.2.1 Methods

The potential indicators used for the development of mental health phenotypes were all items contained in the HADS questionnaire (Zigmond and Snaith, 1983). This scale has been described briefly in previous chapters (*Chapter 6* and *Chapter 7*). The purpose of this scale is to identify participants who are likely to be affected by symptoms of anxiety or depression or both. This questionnaire contains 14 items, of which seven items are related to anxiety symptoms and seven items are related to depression symptoms, and each item has four potential responses that reflect various severities for that item. Each response to each item is scored zero, one, two or three, responses with more severity are scored a higher number. The items contained in the HADS questionnaire may be positive or negative, for example, one item is presented as 'I feel tense or 'wound up' ', whilst another is presented as 'I look forward with enjoyment to things'. The scores for each item were re-ordered as necessary so that responses indicating less anxiety or depression had a lower score.

The final scores range from 0-21 for anxiety and 0-21 for depression, and a score closer to 21 represents more likely to be anxious/ depressed, and a score closer to zero represents less likely to be anxious/ depressed. A full version of the HADS questionnaire is included in *Appendix I*.

For this part of the project, similar to the AUSCAN indicators in the hand phenotype development, the individual items (indicators) rather than the overall scores for anxiety and

depression were used. Each indicator response was dichotomised into lower probability of being affected by that symptom (response of zero or one), and a higher probability of being affected by that symptom (response of two or three).

The population used for the mental health states development were the same individuals used in *Chapter 4*, i.e. all the participants who responded at all three follow-up time points. Again, in line with the analysis population used in *Chapter 4*, individuals who had missing data on more than half of the indicators at each time point were excluded from the analysis.

The process used to identify and remove any indicators that were not necessary in the final states was identical to that employed in *Chapter 4*. The stages are described briefly below:

1. LTA was performed for all of the variables in the model (so initially 14) to investigate what number of states were optimal, based on BIC, entropy and sample size;
2. For the optimal model, each indicator was removed from the model in turn (so initially there were 14 variants of that model) and model fit (BIC/ entropy, defined in *sections 4.6.4.1* and *4.6.4.2*) compared between these models;
3. Sample size of the states was checked to ensure no state had a small size (approximately <5% of participants in that state).
4. The indicator which best improved the model was removed (biggest reduction in BIC/ increase in entropy);
5. Steps 1 to 4 were repeated until removing further terms provided no further improvement to the model (assessed by looking at the interpretation of the states when taking out a further indicator).

As with the model development completed in *Chapter 4*, no restrictions were placed on which indicators could be removed, so indicators could be removed from either anxiety or

depression scales in any order. Once no further indicators were required to be removed to improve the model, the resulting model was then assessed to explore whether it was possible to restrict the item-response probabilities over time and therefore allow state interpretation to be the same at each time point. This process was explained in more detail in *section 5.5*, but briefly, if the restricted model had two of the following then restrictions were enforced: lower BIC, little difference in item-response probabilities within each state at each time point, or non-significant LRT.

9.2.2 Results

5,388 participants responded at all three time points and provided information on at least half of the 14 HADS indicators at each time point. The modelling process has been condensed to display the summary of each removal stage and is displayed in *Table 9.2.1*; the full details of the results from each stage are included in *Appendix H*.

The optimal number of indicators to include in the mental health states was nine items (four anxiety and one depression indicator removed), with four states. While removing further indicators from the model did lead to a substantial difference in BIC, any potential removal of further relevant indicators did not improve the interpretation of the states. The optimum model had a low sample size of 3.5% at baseline, but that state size was 4.8% and 4.9% at the other two time points and was regarded close enough to the 5% boundary criterion. In addition to this, a model with four states was better defined (in terms of state interpretation) than a model with three states, so this model was considered (and as a result, endorsed) as the optimal model.

Table 9.2.1: Summary table of the stages of model development for the mental health states.

Removal stage	Number of indicators	Indicator removed	Number of states	BIC after removal	Entropy after removal	Smallest sample size state
1	14	-	4	123,224	0.881	7.9%
2	13	Feel restless as if I have to be on the move (A)	4	105,954	0.883	7.6%
3	12	Feel as if I am slowed down (D)	4	90,380	0.891	6.0%
4	11	Worrying thoughts go through my mind (A)	4	79,620	0.883	4.3%
5	10	Frightened feeling as if something awful is about to happen (A)	4	68,707	0.874	4.4%
6	**9**	**Can sit at ease and feel relaxed (A)**	**4**	**60,444**	**0.878**	**3.5%**
7	8	Feel tense or 'wound up' (A)	4	50,790	0.881	3.5%

Footnote: BIC= Bayesian Information Criteria; A= Anxiety-related indicator; D= Depression-related indicator; Bold line represents the optimum mental health model.

The optimal model was then assessed to investigate whether restricting item-response probabilities over time could be enforced to maintain the same state interpretation at each time point. The summary of information can be seen in *Table 9.2.2*, along with the associated LRT.

The LRT test produced a *p*-value of <0.001, which implies that restrictions could have been of significant harm to the model. However, considering the interpretational benefit of restricting item-response probabilities, in addition to the fact the BIC's were similar (slightly lower) in the restricted model, and the item-response probabilities were similar in the unrestricted and restricted models at each time point (*Appendix H*) yielding similar phenotype definitions, the restriction was enforced.

Table 9.2.2: Summary of information for the optimum 'null' and 'restricted' model.

	Null model	Restricted model
BIC	60,444	60,046
Entropy	0.878	0.874
Log-likelihood	-29,642.090	-29,752.520
Degrees of freedom	135	63

Footnote: BIC= Bayesian Information Criteria.

The final mental health model is displayed in *Table 9.2.3* and displays the latent state proportions, restricted item-response probabilities and transition probabilities for baseline to 3 years, and 3 years to 6 years.

Firstly, a large proportion of the population were estimated to belong to the first state at all three time points (δ range 0.71-0.74), with approximately 13% in the second state at each time point. The proportion of individuals in the third state decreased over the 6 years (12% to 8%). The fourth state had a consistently low proportion of individuals across the study period, containing 4.4% to 4.7% of the population.

The item-response probabilities were used to derive the state label, as the profile of these estimates reflected the likely probability of endorsing each item by individuals allocated to that state. Individuals allocated to the first state had low probabilities for each of the indicators. At baseline 7.6% of the people allocated to this state reported at least one anxiety characteristic and 5.9% reported a depression characteristic. The state was labelled 'no anxiety/ depression'.

Table 9.2.3: LTA estimates of the mental health phenotypes with restricted item-response probabilities.

Mental health restricted model (n= 5,388)	No anxiety/ depression	Mild depression	High anxiety	Anxiety & depression
Latent state proportions				
Baseline	0.710	0.129	0.117	0.044
3 years	0.734	0.129	0.093	0.044
6 years	0.742	0.135	0.076	0.047
Item-response probabilities				
Feel tense or 'wound up'	0.036	0.234	0.582	0.872
Still enjoy the things I used to enjoy	0.013	0.406	0.096	0.764
Can laugh and see the funny side of things	0.003	0.115	0.048	0.521
Feel cheerful	0.001	0.067	0.054	0.503
Frightened feeling like 'butterflies' in my stomach	0.006	0.029	0.434	0.604
Lost interest in my appearance	0.020	0.218	0.103	0.450
Look forward with enjoyment to things	0.010	0.392	0.147	0.860
Get sudden feelings of panic	0.009	0.043	0.653	0.710
Can enjoy a good book or radio or television programme	0.007	0.074	0.039	0.255
Latent transition probabilities				
Baseline to 3 years	No anx/ dep	Mild dep	High anx	Anx & dep
No anxiety/ depression	**0.947**	0.031	0.017	0.005
Mild depression	0.243	**0.696**	0.015	0.046
High anxiety	0.255	0.035	**0.646**	0.065
Anxiety & depression	0.017	0.308	0.067	**0.608**
3 years to 6 years	No anx/ dep	Mild dep	High anx	Anx & dep
No anxiety/ depression	**0.946**	0.035	0.018	0.002
Mild depression	0.149	**0.764**	0.000	0.087
High anxiety	0.284	0.000	**0.640**	0.076
Anxiety & depression	0.057	0.239	0.088	**0.617**

Footnote: No anx/ dep= 'no anxiety or depression'; Mild dep= 'mild depression'; High anx= 'high anxiety'; Anx& dep= 'anxiety and depression'; Bold entries represent stability between time points.

The second state had higher probabilities for indicators that were more related to depression than anxiety. For example, unable to ‘still enjoy the things I used to enjoy’, unable to ‘look forward with enjoyment to things’ and ‘lost interest in my appearance’ were more probable than the other indicators (with the exception of ‘feel tense or ‘wound up’ ’). At baseline 38.7% of the people allocated to this state reported at least one anxiety characteristic and 94.7% reported a depression characteristic. The second state was therefore labelled ‘mild depression’ as the individuals in this state reflected participants with a modest increase in probability of symptoms of depression (50% of individuals in this state reported only one symptom of depression, 29% reported two symptoms, while 14% reported three or more).

The third state reflected individuals who were more likely to respond negatively to the indicators ‘feel tense or ‘wound up’ ’, ‘frightened feeling like ‘butterflies’ in my stomach’ and ‘get sudden feelings of panic’ which are anxiety related items from the HADS. At baseline 96.6% of the people allocated to this state reported at least one anxiety characteristic and 47.1% reported a depression characteristic. The profile of this state was more decisive than participants in the second state, in combination with larger item-response probabilities, the third state was labelled ‘high anxiety’ (31% of individuals in this state reported one anxiety symptom, 37% reported two symptoms, and 28% reported all three).

The probabilities in the fourth state were all larger than any of those in the previous three states for each indicator. As a result, these probabilities reflected a (small) sub-sample of the population who had evidence of both anxiety and depression symptoms. At baseline 98.3% of the people allocated to this state reported at least one anxiety characteristic and 100% reported a depression characteristic, and therefore the state was labelled ‘anxiety and depression’.

Exploring the transition probabilities of the 3 year periods between baseline to 3 years and 3 years to 6 years firstly showed that participants with little or no symptoms of anxiety and depression were more likely to also have minimal symptoms 3 years later as 95% of the 'no anxiety/ depression' state remained in that state 3 years later (τ= 0.947/ τ= 0.946). Similarly, the other three states showed stability with a high probability remaining in the same state (all τ>0.608). For the 'mild depression' and 'high anxiety' states, transition probabilities into different states were low with the exception of transitioning into the 'no anxiety/ depression' state (τ range 0.15 to 0.28 over both time periods), however this was more likely for individuals in the 'high anxiety' state. This result highlighted in our sample, that participants with either symptoms of depression or of anxiety were unlikely to develop symptoms of the other condition and therefore did not transition into the 'anxiety and depression' state (all τ<0.087).

Whilst there was a high probability of individuals in the most severe state 'anxiety and depression' remaining in the same state 3 years later (τ>0.608), the probability of individuals transitioning into the 'mild depression' state 3 years later (showing an apparent improvement in their anxiety symptoms in general) was considerable (τ= 0.31 for baseline to 3 years and τ= 0.24 for 3 years to 6 years). The probability of this group moving into 'high anxiety' or 'no anxiety/ depression' was low. It is important to note that the number of individuals estimated to be in the more severe state at each time point was smaller than other states (approximately 5%) so these probabilities were likely to be less accurate due to the small sample size of this state. However, this finding highlighted that individuals in this population appeared more likely to experience an improvement in their anxiety symptoms 3 years later, if they had both anxiety and depression symptoms to start with (this coincides with the larger probability of moving from 'high anxiety' to 'no anxiety/ depression' than from 'mild depression' to 'no anxiety/ depression').

9.2.3 Summary

This section developed a latent variable of mental health states that categorised individuals into one of four states at each of baseline, 3 years and 6 years, which were labelled ‘no anxiety/ depression’, ‘mild depression’, ‘high anxiety’ and ‘anxiety and depression’. At least two-thirds of individuals in each state were estimated to remain in the same state 3 years later. In addition to this, the transition probabilities appeared to highlight a group with apparent improvement in anxiety symptoms for the participants in the most severe mental health state between time points, with a smaller estimated number showing improvement of depression symptoms.

The prevalence of depression and anxiety symptoms were not explicitly assessed in this section. However, it can be estimated that at baseline approximately 17.3% of the population had some level of depression symptoms (the estimated combined proportion in the ‘mild depression’ and ‘anxiety and depression’ states), while approximately 16.1% of the population had some anxiety symptoms (the estimated combined proportion in the ‘high anxiety’ and ‘anxiety and depression’ states). Point prevalence estimates for symptoms of depression in older adults (aged >65 years) are estimated between eight to 16% (Blazer, 2003; Park and Unützer, 2011), while prevalence rates of anxiety tend to be lower, in regions between four and 10% (Wolitzky-Taylor et al., 2010; Lenze and Wetherell, 2011). However a study of older Irish adults found higher rates of anxiety symptoms compared to depression symptoms (TILDA study, Barrett et al., 2011). The prevalence found in this study for depression and anxiety was larger, however these estimates were likely to be an overestimation as they included individuals with mild symptoms, as well as containing a majority of individuals aged 50 to 65 years where symptoms are more common compared to older age groups (Fiske et al., 2009; Rodda et al., 2011).

At 6 years follow-up these estimates had increased for depression symptoms to 18.2%, while anxiety symptoms had reduced to 12.3%. In addition to this, the transition probabilities implied that participants with anxiety symptoms were more likely to see a reduction in their condition. Other research has found that the prevalence of anxiety-related issues generally decline into older age, however some aspects of anxiety can increase (such as fear of falling) (Lenze and Wetherell, 2011). Alternatively, in previous research the prognosis for those with depression is not as positive; the majority of individuals with depressive symptoms do not recover, whilst also being at an increased risk of mortality (Rodda et al., 2011; Chesney et al., 2014). The course of anxiety and depression symptoms found in this section have some similarities with previous literature.

Strengths and limitations of this process are similar to those of the latent hand state development. Using LTA to develop states of mental health permitted the development of sub-groups to represent the most common profiles of mental health in our population. Developing new latent states instead of using the pre-existing 'cut-offs' for the HADS also created more opportunity for different patterns of improvement/ deterioration that is not permitted with the original categorical scale. It could have been possible to use the original cut-offs for the HADS questionnaire to explore the longitudinal pattern of mental health (by creating an observed matrix of transitional probabilities generated by simple frequencies/ proportions of the HADS categories). However, using LTA permitted the flexibility of creating the most relevant states for the population instead of using pre-set cut-off points, and has also removed indicators that, statistically, did not distinguish between mental health states.

Limitations include some individuals being excluded due to missing data, the potential loss of information as a result of dichotomisation, and subjective labelling of the mental health states. Only a small number of individuals were deleted due to missing data (363/ 5,751

(6.3%)), and the removal of individuals were consistent with the approach taken in *Chapter 4*. Dichotomisation was enforced to make the computation process quicker and to ease interpretation. Unlike the development of the hand states, the labelling of the mental health states were not developed in consultation with an 'expert'. However, states were labelled logically and were agreed within the supervisory team.

The next section uses the ALTA technique to investigate the longitudinal relationship between the latent states of mental health and the latent hand states developed in *Chapter 4*.

9.3 Associative Latent Transition Analysis (ALTA) - Overview

The statistical technique used in the remainder of this chapter is Associative Latent Transition Analysis (ALTA, briefly explained in *section 3.9.6*), which is a further extension to LTA. The main purpose of ALTA is to explore whether two latent variables (health domains, for example) are associated with each other (either cross-sectionally and/ or longitudinally), and where the strongest associations occur. The technique is described below in detail based on the description of the original ALTA paper (Flaherty, 2008), but using the two domains that are investigated in the rest of the chapter (hand states developed in *Chapter 4*, and mental health states developed in *section 9.2*).

In the ALTA process there needs to be a dependent latent variable which is the primary focus of the research, which in this chapter is the mental health states, and a predictor latent variable (here hand states), and therefore changes in membership in the predictor variable could potentially result in changes to membership in the dependent variable. The main research question is: does a change in hand state lead to a change in mental health state membership?

All of the estimated probabilities produced in the ALTA technique are represented in three main sets of estimates; the first are the β (beta) estimates which represent the probability distribution of the mental health states at time 1 (T1), conditional on hand state at T1. The second sets of values are defined as the ε (epsilon) estimates. These denote the probability of hand state at time 2 (T2), conditional on latent class memberships for both hand and mental health states at T1. The final sets of results from the ALTA are represented by the η (eta) estimates which represent the probabilities of mental health state membership at T2, conditional on previous and concurrent levels of hand state membership (T1 and T2) and the previous mental health state membership (T1). This final set of estimates are the most relevant to address the research question, however the prior two sets of estimates are relevant to further understand how the states change over time.

To fully utilise the ALTA technique, four separate models are computed that incorporate the various levels of association between the two domains (mental health and hand pain/ problems). All four models can be run independently of each other (so the results of one model are not required to compute another). The first model, labelled the 'Independence' model restricts each of the ALTA parameters (β's, ε's and η's) to be equal in each instance (e.g. all β's are the same, all ε's are the same (potentially different to the β's) etc.). Therefore, the main assumption of this model is no relationship between the domains at any time point. The restriction of the β's ensures that the probability of mental health state membership does not depend on hand state membership at the same time point. The restriction of the ε's ensures that the probability of the hand state membership at T2 relates only to hand state membership at T1 and does not depend on mental health state membership at T1. The restriction of the η's ensures that T2 mental health state membership only depends on T1 mental health state, and does not depend on T1 or T2 hand state membership. This model is similar to completing two separate LTA estimations

where no relationship between hand and mental health states are considered. This relationship would be unlikely in this scenario considering the baseline characteristics displayed in *Table 6.2.1*, as individuals in more severe states had higher scores of anxiety and depression symptoms.

The second model is labelled the 'cross-sectional' model which applies restrictions on the longitudinal ALTA parameters (ε's and η's), but freely estimates the parameters that represent the cross-sectional relationship between the domains (β's). Therefore, the probability of mental health state at T1 given hand state membership at T1 were freely estimated, but all probabilities that investigated any longitudinal changes in state memberships were restricted to be equal (by restricting the ε's and the η's). In more simple terms, mental health state memberships were permitted to depend on concurrent hand state memberships but no longitudinal relationship is permitted. This model would be suitable considering previous cross-sectional association from *section 6.2*, but could potentially be too restrictive for this research question given the possibility of a longitudinal association.

The third model is labelled the 'longitudinal' model which enforces equality restrictions on the β's, but permits the ε's and η's to be freely estimated. The model represents a scenario where cross-sectional relationships at T1 between mental health state and hand state are not permitted (the β's), but an individual's hand state membership at T2 can depend on previous mental health and hand state membership (the ε's), and mental health state membership at T2 can be influenced by both previous mental health state membership and previous and concurrent hand state membership (the η's). Similar to the independence model, this scenario may be unlikely because a cross-sectional relationship between hand state and mental health was apparent from the results in *Chapter 6*.

The fourth model, labelled the 'full association' model contains no restrictions on any of the ALTA parameters, and therefore the two domains are associated cross-sectionally and longitudinally. Therefore, in this study, an individual's mental health state membership at T2 can depend on their previous mental health state membership, and concurrent and previous hand state membership. A cross-sectional relationship between hand state and mental health could be presumed from prior findings (subject to further evidence), so the full association model could be regarded as the most informative in this analysis to assess whether a longitudinal relationship between the two domains exists in this population.

As described in the original paper (Flaherty, 2008), all four models should be computed and then relevant comparisons made to investigate which model is the most appropriate. Due to the restriction enforcements of these four models, they can be considered as nested models, and therefore comparison between models on the difference in likelihoods (G^2) and degrees of freedom can help determine the best fitting model (Agresti, 1990). Each of the four models should replicate their own respective largest log-likelihood estimates to ensure the estimates are robust and accurate, identical to the process of simulating LTA models previously.

The ALTA process is a powerful technique that can highlight longitudinal associations between related health domains. However, there are some current drawbacks. The technique is a complex longitudinal process, particularly when a large number of parameters are being estimated. As a result, the availability of software to perform ALTA is limited. Researchers at Penn State University (Bray et al., 2010) and the University of New Mexico (Witkiewitz and Villarroel, 2009) have started to implement almost parallel techniques in Mplus. However, due to limitations of the Mplus software, and lack of readily available support, full implementation of ALTA models is not currently available in Mplus.

Flaherty has developed software to implement the ALTA technique (which was used to produce the estimates in his 2008 paper) which is readily available. However, additional issues arise concerning the operating system the ALTA software is run on (as it was originally written in the Linux operating system). Further work (aside from this thesis) is required to ensure the technique is more readily available for future researchers.

9.4 ALTA - Methods

The ALTA technique described in *section 9.3* was applied in order to model the relationship between hand state and mental health state over a 6 year period. This analysis included all three time points, baseline, 3 years and 6 years. Of the three identified published papers using ALTA, only one (Witkiewitz and Villarroel, 2009) has considered a third time point in the analysis. However, the analysis in this publication contained various combinations of time points due to alternative recruitment times for their domains (alcohol data collected at baseline, 3, 6, 9 and 12 months, and psychological state at baseline, 6 and 12 months). The authors assessed alcohol use at 6 months based on current psychological state (6 months), previous alcohol use (3 months) and previous psychological state (baseline) (with a similar analysis for alcohol use at 12 months). The analysis presented here will describe the relationships between T1 (baseline survey) and T2 (3 years follow-up), and then those between T2 and T3 (6 years follow-up) (estimated in the same computation). This was similar to the previous LTA analyses; while the probabilities were estimated in the same process, the relationships between T2 and T3 did not take into account state memberships at T1.

Numerous issues were encountered when performing the ALTA which are reported in full in the results (*section 9.5*). However, in order to perform a successful computation of the ALTA parameters, individuals with missing data had to be excluded. Therefore, only

individuals with responses for all hand and mental health indicators (eight indicators for hand, nine indicators for mental health) at each of the follow-up points were included in the analysis. The completed analysis was performed in the ALTA software using Ubuntu Linux 14.04 LTS.

9.5 ALTA - Results

9.5.1 Simulation

4,903 out of the 5,617 participants (87%) included in the analysis of hand states (*Chapters 4* to *6*) responded at each time point, and provided responses to all hand indicators and mental health indicators at baseline, 3 years and 6 years, and were therefore included in the ALTA analysis. The separate LTA model specifications for hand state (displayed in *Table 6.2.1*) and mental health (displayed in *Table 9.2.3*) were considered for the ALTA technique. Therefore, the independent domain (hand state) had five latent states at each of the three time points with item-response probabilities constrained to be the same for each time point (and therefore state definition remained identical at each follow-up time point). The dependent domain (mental health state) had four latent states at each of the three time points with the same item-response constraint.

All four of the ALTA models (independent, cross-sectional, longitudinal and full association) were initially attempted in the Mplus software. However, enforcing parameter restrictions in the first three models created complications and therefore, were not computed successfully and simulations were not completed. Despite seeking assistance from Mplus support, sufficient progress was not made and unfortunately models for the independent, cross-sectional and longitudinal models were not computed. The full association model was also attempted in Mplus, but again, due to limitations of the

software (such as Mplus struggling to identify a model with higher order interaction terms, more details mentioned in Witkiewitz and Villarroel, 2009), certain aspects of the full association model were not achievable in Mplus, and again, a computable model was not achieved.

However, through communication with Dr. Flaherty, University of Washington, the ALTA software he developed was received, but this brought an additional complication. Dr. Flaherty was only available for a short period of time, institutional (University of Washington) support was limited and a version of the ALTA software to run in the Windows software was not readily available (at least not suitable for the PhD project timeline). Therefore, the only suitable version of the ALTA software was a version to run in the Linux operating system. The Linux system is a powerful operating system, however, a substantial amount of time would have been required to become a competent user. As a result of this potentially steep learning curve, Dr. Flaherty assisted the project by running three versions of the full association model on an anonymised version of the dataset on his personal system, and then emailed over the results. Assistance from Dr. Flaherty was limited to running this model according to my (the candidate's) specifications and all interpretations of the model are my own.

Of the three versions of the simulated model, two were given random start values to begin computation, while the third used the item-response probabilities generated from the Mplus output for the hand and mental health states (both of which have been reported previously) as start values. Each of the models took around 20 hours to simulate and all reported different likelihood results ('40,531.82', '39,974.90' and '40,032.61'), implying that the model was not easily estimable, and the model estimates should be viewed with caution. However, the model based on the previously derived item-response probabilities and one of the models initiated with random start values did have similar log-likelihoods

('39,974.90' and '40,032.61'), and therefore the model initiated with item-response probabilities is reported here. The other two models gave similar, albeit not identical, estimates to the model presented here, and similar interpretations could be drawn from those models and the associations presented below. In all of the following tables (*Table 9.5.1* to *Table 9.5.10*), estimates in bold represent individuals estimated to remain in the same state(s) as the prior time point (where applicable).

9.5.2 Concurrent relationship of hand and mental health states

The β estimates represent the probability of concurrent relationship between hand states and mental health. The 60 probabilities generated by this stage of analysis (five hand states × four mental health states × three time points) were split into three tables, one for each time point. The two tables that assess the relationships at baseline and 3 years are displayed in *Table 9.5.1* and *Table 9.5.2* (the table for 6 years, which had similar results to 3 years, is displayed in *Appendix J*, where all the ALTA tables are presented).

At baseline, a large probability of individuals estimated to be in the 'least affected' state were also estimated to be in the 'no anxiety/ depression' state (β= 0.852), while the majority of individuals in other baseline hand states had a low probability of being in the 'no anxiety/ depression' state (β's<0.170). Of those in the 'high pain' state at baseline, the majority were estimated to be in 'mild depression' state at baseline (β= 0.714), whilst there were similar probabilities for those in the 'poor gross function' state being in the 'high anxiety' state (β= 0.719). Individuals in the more severe hand states, 'high pain and poor gross function' and 'severely affected', were most likely to be classified in the 'anxiety and depression' state (β= 0.491 and β= 0.416, respectively), almost three times as likely as being classified in 'no anxiety/ depression' (β= 0.169 and β= 0.157, respectively).

Table 9.5.1: Time 1 β estimates for concurrent relationships; baseline mental health state dependent on baseline hand state.

Baseline hand / Baseline mental health	No anxiety/ depression	Mild depression	High anxiety	Anxiety & depression
Least affected	0.852	0.050	0.048	0.050
High pain	0.170	0.714	0.039	0.078
Poor gross function	0.107	0.132	0.719	0.042
High pain & poor gross function	0.169	0.172	0.169	0.491
Severely affected	0.157	0.154	0.273	0.416

Table 9.5.2: Time 2 β estimates for concurrent relationships; 3 year mental health state dependent on 3 year hand state.

3 year hand / 3 year mental health	No anxiety/ depression	Mild depression	High anxiety	Anxiety & depression
Least affected	0.451	0.166	0.149	0.235
High pain	0.340	0.343	0.172	0.145
Poor gross function	0.249	0.317	0.311	0.124
High pain & poor gross function	0.121	0.328	0.257	0.294
Severely affected	0.269	0.177	0.131	0.422

At 3 years the estimated probabilities changed, in particular, none of the estimates were greater than 0.5, implying the cross-sectional patterns were not highly similar to those demonstrated at baseline. Individuals in 'least affected' at 3 years were still more likely to be in 'no anxiety/ depression' state at 3 years (β= 0.451) than any of the other mental health states, however a quarter of individuals in 'least affected' were also estimated to be in 'anxiety and depression'. Participants in 'high pain' at 3 years were most likely to be classified in 'no anxiety/ depression' or 'mild depression' at 3 years (β= 0.340 and β= 0.343, respectively) while those in 'poor gross function' were spread mostly amongst the least severe mental health states (β= 0.249 for 'no anxiety/ depression, β= 0.317 for 'mild

depression' and β= 0.311 for 'high anxiety'). Therefore, the findings from baseline ('high pain' related with 'mild depression' and 'poor gross function' related with 'high anxiety') were not consistent at 3 years. Those in 'high pain and poor gross function' at 3 years were estimated to be most likely in one of the three more severe mental health states (β= 0.328 for 'mild depression', β= 0.257 for 'high anxiety' and β= 0.294 for 'anxiety and depression') while those in 'severely affected' were most likely to be in either 'anxiety and depression' (β= 0.422) or 'no anxiety/ depression' (β= 0.269) at 3 years.

The results given in these tables have identified that while some strong relationships between hand state and mental health were visible at baseline, these were less apparent at 3 years (and at 6 years, *Appendix J*). The subsequent analyses explores whether there were longitudinal associations between these two domains.

9.5.3 Association of hand states with prior hand and mental health states

Even though the main interest of the ALTA technique was the mental health outcome at 3 years (and 6 years), which is presented in the next section (*section 9.5.4*), the process also generates an intermediate set of probabilities, the ε estimates which assess the follow-up hand state, dependent on previous hand state and previous mental health state. While these estimates were not the central focus of this research question, investigating these estimates were important to understand how the predictor state (hand states) changed over time, and also ensuring interpretations made in the next section (*9.5.4*) were not made out of context (by ignoring the previous pattern of estimates).

Due to the number of latent states in each of the domains, this stage of analysis produces 200 probabilities (five hand states at time n × five hand states at time n-1 × four mental health states at time n-1 × two time period assessments (baseline to 3 years, and 3 years to

6 years)), that were sectioned into five tables for each time period (baseline to 3 years, and 3 years to 6 years), therefore, 10 tables in total. Only a small subset of tables are displayed in this chapter while the remaining tables are presented in *Appendix J*. The main interest in this sub-section was individuals who were in the 'least affected' hand state, and those who were in the most severe hand state ('severely affected') at the previous time point, representing longitudinal changes in those at the different endpoints of the spectrum of hand pain severity.

9.5.3.1 'Least affected' at baseline

Table 9.5.3 displays the probability (ε) of hand state membership at 3 years, dependent on baseline mental health state within individuals that were estimated to be in the 'least affected' hand state at baseline. The column in bold highlights the individuals who were estimated to remain in the 'least affected' hand state at 3 years. The probabilities of individuals remaining in this state have previously been shown to be high (τ= 0.867, *Table 6.3.1*), and the estimates in *Table 9.5.3* were also high (all ε>0.69). Considering baseline mental health, those in 'no anxiety/ depression' were most likely to remain in 'least affected' at 3 years (ε= 0.94), compared to those with 'high anxiety' or 'anxiety and depression' (ε= 0.79 and ε= 0.69, respectively); however, these probabilities were still relatively large (ε>0.5). Of the individuals who were most likely to transition from 'least affected' at 3 years, 'high pain' was the most probable state, which was much more likely for those in the 'high anxiety' or 'anxiety and depression' states at baseline (ε= 0.16 and ε= 0.21, respectively) than for those in 'no anxiety/ depression' or 'mild depression' (ε= 0.04 and ε= 0.08, respectively). Probabilities of hand state membership at 6 years for individuals in 'least affected' state at 3 years were similar to the baseline results (*Table 9.5.4*).

Table 9.5.3: Time 2 ε estimates reflecting probability of change in hand state; 3 year hand state for those in 'least affected' hand state at baseline, dependent on baseline mental health state.

3 year hand BL mental health	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
No anxiety/ depression	**0.935**	0.039	0.019	0.007	0.001
Mild depression	**0.875**	0.077	0.013	0.029	0.006
High anxiety	**0.785**	0.155	0.044	0.017	0.000*
Anxiety & depression	**0.692**	0.213	0.054	0.035	0.006

Footnote: *: probabilities of exactly '0.00' and '1.00' should be interpreted with caution (potential small sample size).

Table 9.5.4: Time 3 ε estimates reflecting change in hand state; 6 year hand state for those in 'least affected' hand state at 3 years, dependent on 3 year mental health state.

6 year hand 3 year mental health	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
No anxiety/ depression	**0.910**	0.052	0.025	0.011	0.002
Mild depression	**0.837**	0.095	0.050	0.018	0.000*
High anxiety	**0.854**	0.083	0.023	0.041	0.000*
Anxiety & depression	**0.726**	0.173	0.064	0.031	0.007

Footnote: *: probabilities of exactly '0.00' and '1.00' should be interpreted with caution (potential small sample size).

9.5.3.2 'Severely affected' at baseline

Table 9.5.5 displays the ε estimates of individuals in 'severely affected' state at baseline. Individuals in the 'severely affected' state at baseline had a slightly higher probability of remaining in 'severely affected' at 3 years if they were in the 'anxiety and depression' state compared to the 'no anxiety/ depression' state at baseline (ε= 0.52 vs. ε= 0.41) (*Table 9.5.5*). Distinct patterns of transitions from 'severely affected' state at baseline into less severe states at 3 years were difficult to identify. However, investigating the next time period (3 years to 6 years) identified a larger difference in those likely to remain in 'severely affected' at 6 years (from 3 years), if they were in 'anxiety and depression' at 3 years, compared to those in 'no anxiety/ depression' at 3 years (ε= 0.41 vs. ε= 0.09) (*Table 9.5.6*).

Table 9.5.5: Time 2 ε estimates reflecting change in hand state; 3 year hand state for those in 'severely affected' hand state at baseline, dependent on baseline mental health state.

3 year hand BL mental health	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
No anxiety/ depression	0.000*	0.000*	0.194	0.400	**0.406**
Mild depression	0.168	0.236	0.194	0.116	**0.287**
High anxiety	0.036	0.123	0.274	0.174	**0.393**
Anxiety & depression	0.023	0.060	0.280	0.122	**0.515**

Footnote: *: probabilities of exactly '0.00' and '1.00' should be interpreted with caution (potential small sample size).

Table 9.5.6: Time 3 ε estimates; 6 year hand state for those in 'severely affected' hand state at 3 years, dependent on 3 year mental health state.

6 year hand 3 year mental health	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
No anxiety/ depression	0.136	0.279	0.315	0.181	**0.090**
Mild depression	0.087	0.000	0.253	0.423	**0.237**
High anxiety	0.062	0.142	0.315	0.000*	**0.480**
Anxiety & depression	0.060	0.000*	0.252	0.278	**0.410**

Footnote: *: probabilities of exactly '0.00' and '1.00' should be interpreted with caution (potential small sample size).

In addition to this, individuals in 'no anxiety/ depression' at 3 years were more likely to move to a less severe hand state at 6 years ('least affected' ε= 0.14, 'high pain' ε= 0.28, 'poor gross function' ε= 0.32), compared to those in 'anxiety and depression' at 3 years ('least affected' ε= 0.06, 'high pain' ε= 0.00, 'poor gross function' ε= 0.25).

9.5.4 Association of mental health states with concurrent hand state and prior mental health and hand states

The final sets of estimates produced in the ALTA process were the η estimates which assess the follow-up mental health state, dependent on concurrent and previous hand state membership, and previous mental health state membership, hence these estimates were most relevant for the research question formulated in *section 9.3*: does a change in hand state lead to a change in mental health state membership? With the number of latent states

in each of the domain, this step of the analysis produced 800 probabilities (four mental health states at time n × five hand states at time n × four mental health states at time n-1 × five hand states at time n-1 × two time period assessed (baseline to 3 years, and 3 years to 6 years)), that are sectioned into 20 tables for each time period (baseline to 3 years, and 3 years to 6 years), with therefore 40 tables in total. As a result of the number of parameters estimated, some of the combinations of latent states were subject to small sample sizes, and should be interpreted with caution.

Similar to *9.5.3*, only a subset of tables are displayed in this chapter but the additional tables are presented in *Appendix J*. Again, similar to the previous section, the results focus on individuals that were in the 'least affected' hand state, and those in the most severe hand state ('severely affected').

9.5.4.1 'Least affected' at baseline

Table 9.5.7 displays the η estimates for 3 year mental health state dependent on 3 year hand state membership, given classification in 'least affected' hand state and 'no anxiety/ depression' at baseline. Firstly, individuals that remained in 'least affected' at 3 years were highly likely to remain in the 'no anxiety/ depression' state at 3 years (η= 0.91), whilst only a half of those who transitioned to the worst hand state 'severely affected' were estimated to remain in 'no anxiety/ depression' at 3 years (η= 0.52). However, of the individuals who transitioned from 'least affected' hand state to 'high pain and poor gross function' or 'severely affected' at 3 years, a third to a quarter were estimated to transition into the 'anxiety and depression' mental health state at 3 years (η= 0.38 and η= 0.23, respectively). These probabilities were similar for the 3 year to 6 year period (*Appendix J*).

Table 9.5.7: Time 2 η estimates; 3 year mental health state, dependent on 3 year hand state membership, given membership in 'least affected' hand state and 'no anxiety/ depression' at baseline.

3 year mental health 3 year hand	No anxiety/ depression	Mild depression	High anxiety	Anxiety & depression
Least affected	**0.909**	0.044	0.017	0.030
High pain	0.180	0.675	0.059	0.086
Poor gross function	0.022	0.117	0.778	0.084
High pain & poor gross function	0.155	0.290	0.179	0.377
Severely affected	0.523	0.249	0.000*	0.228

Footnote: *: probabilities of exactly '0.00' and '1.00' should be interpreted with caution (potential small sample size).

Of those estimated to be in the 'least affected' hand state and 'no anxiety/ depression' mental health state at baseline, those who transitioned into the 'high pain' state at 3 years were estimated to have a high probability of also transitioning into the 'mild depression' hand state at 3 years (η= 0.68), which was similar in the 3 years to 6 years analysis (η= 0.69, *Appendix J*). Those who transitioned into 'poor gross function' at 3 years (from 'least affected' and 'no anxiety/ depression' at baseline) had a high probability of transitioning into 'high anxiety' at 3 years (η= 0.78), which was also similar in the 3 years to 6 years analysis (η= 0.67, *Appendix J*). These relationships were similar in the concurrent analysis at baseline (*Table 9.5.1*), but not at follow-up points.

Investigating those in 'least affected' and in 'anxiety and depression' at baseline, the probability of those estimated to remain in 'anxiety and depression' at 3 years, if remained in 'least affected' at 3 years, was relatively large (η= 0.61) (*Table 9.5.8*), with a similar probability in the 3 year to 6 year time period (η= 0.50) (*Appendix J*). Of the individuals who transitioned from 'least affected' to 'severely affected' at 3 years, the probability of remaining in 'anxiety and depression' was large (η= 1.00), while those who transitioned into 'high pain and poor gross function' were more likely to be in 'high anxiety' (η= 0.66).

However, these estimates appeared sensitive to the small sample size, and as a result the probabilities of this pattern were not consistent in the 3 year to 6 year period (*Appendix J*).

Table 9.5.8: Time 2 η estimates; 3 year mental health state, dependent on 3 year hand state membership, given membership in 'least affected' hand state and 'anxiety and depression' at baseline.

3 year mental health 3 year hand	No anxiety/ depression	Mild depression	High anxiety	Anxiety & depression
Least affected	0.166	0.088	0.138	**0.608**
High pain	0.779	0.177	0.018	0.026
Poor gross function	0.122	0.545	0.228	0.105
High pain & poor gross function	0.000*	0.000*	0.659	0.341
Severely affected	0.000*	0.000*	0.000*	1.000*

Footnote: *: probabilities of exactly '0.00' and '1.00' should be interpreted with caution (potential small sample size).

9.5.4.2 'Severely affected' and 'no anxiety/ depression' at baseline

Considering the individuals who were estimated to be in 'severely affected' hand state and in 'no anxiety/ depression' states at baseline, a fifth were estimated to remain in 'no anxiety/ depression' at 3 years, if remained 'severely affected' at 3 years (η= 0.20) (*Table 9.5.9*). Individuals remaining in 'severely affected' hand state at 3 years were more likely to transition into 'anxiety and depression' (η= 0.57) or 'high anxiety' (η= 0.23) at 3 years than stay with 'no anxiety/ depression'. However, these patterns were not demonstrated in the 3 years to 6 years follow-up, potentially due to a small sample size (*Appendix J*).

Table 9.5.9: Time 2 η estimates; 3 year mental health state, dependent on 3 year hand state membership, given membership in 'severely affected' hand state and 'no anxiety/ depression' at baseline.

3 year mental health 3 year hand	No anxiety/ depression	Mild depression	High anxiety	Anxiety & depression
Least affected	0.000*	1.000*	0.000*	0.000*
High pain	0.000*	0.000*	1.000*	0.000*
Poor gross function	0.326	0.310	0.000*	0.365
High pain & poor gross function	0.158	0.000*	0.000*	0.842
Severely affected	**0.204**	0.000*	0.228	0.568

Footnote: *: probabilities of exactly '0.00' and '1.00' should be interpreted with caution (potential small sample size).

9.5.4.3 'Severely affected' and 'anxiety and depression' at baseline

Individuals who were estimated to be in 'severely affected' and 'anxiety and depression' at baseline were highly likely to remain in 'anxiety and depression' at 3 years if also in 'severely affected' at 3 years (η= 0.81) *(Table 9.5.10)*. In addition to this, the probability was higher for the 3 year to 6 year period (η= 0.94) (*Appendix J*). For those in 'severely affected' and 'anxiety and depression' at baseline, participants who were estimated to transition to a less severe hand state at 3 years were also estimated to transition to a less severe mental health state at 3 years, in particular into 'mild depression' (η= 0.46 for those estimated to transition to 'high pain', η= 0.90 for transitioning to 'poor gross function', and η= 0.39 for transitioning to 'high pain and poor gross function', *Table 9.5.10*). Similarly, in general, transitioning to a less severe hand state at 3 years was associated with transitioning to a less severe mental health state at 6 years (*Appendix J*).

Table 9.5.10: Time 2 η estimates; 3 year mental health state, dependent on 3 year hand state membership, given membership in 'severely affected' hand state and 'anxiety and depression' at baseline.

3 year mental health 3 year hand	No anxiety/ depression	Mild depression	High anxiety	Anxiety & depression
Least affected	0.000	0.000	1.000	0.000
High pain	0.143	0.458	0.400	0.000
Poor gross function	0.096	0.904	0.000	0.000
High pain & poor gross function	0.000	0.386	0.208	0.406
Severely affected	0.047	0.047	0.100	**0.807**

9.6 Summary

The ALTA technique is a novel approach to investigating the longitudinal relationship between patterns of health domains. Whilst limitations exist, plausible relationships between hand state and mental health were identified, however, caution in probability estimates is advised as discussed below. An overall theme appeared to be that those who were in the least severe states for both hand and mental health states were more likely to remain in those states at the subsequent follow-up point. In addition to this, participants who were in the more severe hand states and mental health state were also more likely to remain in the more severe hand states and mental health state 3 years later. Also, probable relationships between the 'high pain' and 'mild depression' states, and the 'poor gross function' and 'high anxiety' states, were found over cross-sectional and longitudinal investigations. Probabilities in the 3 year to 6 year period provided similar interpretations to the baseline to 3 year period in the majority of assessments.

A further pattern identified was that improvements from the more severe hand states over the 3 year period were frequently paralleled with an improvement in mental health state. Previous literature has identified an association between musculoskeletal health and depression/ mood (Bair et al., 2003; Demyttenaere et al., 2007; Rosemann et al., 2007;

Thomas et al., 2008; Arola et al., 2010; Kroenke et al., 2011), or anxiety conditions (Mallen et al., 2007; Arola et al., 2010). While these studies have not necessarily focussed on hand conditions, it is conceivable that improvements in pain or function could result in a more positive outlook.

The technique of ALTA has the promise to be a very powerful statistical approach for future research with some modifications and improvements. With regard to this research question, clearly additional work is required to generate a reliable assessment of longitudinal associations between mental health and hand conditions in older people. More specifically, all four of the ALTA models should be computed to determine the overall relationship between the two domains, with each of the four models computed with a repeated and reliable G^2/ log-likelihood estimate. In addition to this, interpretation of findings can be complex given the large number of probability estimates generated.

The main limitation of the ALTA work completed here was that the model described did not have a consistently repeated log-likelihood, therefore, the estimates displayed were not robust and were interpreted with caution. As the log-likelihood in the presented model used the state item-response probabilities from the estimated LTA model, and that model was similar to a model computed with random start values, one can presume that the estimates in a robust model would not be too dissimilar to the values presented. However, replication of these estimates (and log-likelihood) is required.

The stability of the 'severely affected' state was reasonably high in the standard LTA analysis (τ= 0.68 for 'severely affected', *Table 6.3.1*), however, the probabilities for individuals remaining in 'severely affected' in this analysis were lower (ε's ranged 0.29-0.52) when investigating the 'severely affected' state at 3 years in the ALTA analysis (*section 9.5.3.2*, *Table 9.5.5*). Exploring the pathways of individuals in the 'severely

affected' hand states (presented in a flow diagram *Figure K.1* in *Appendix K* estimating the expected number of individuals in the 'severely affected' and 'anxiety and depression' states over 3 years), it was seen that only 76 individuals were estimated to remain in the 'severely affected' state at 3 years if they were in 'anxiety and depression' at baseline (the largest estimated proportion from *Table 9.5.1*). Therefore, due to the reduced sample size when assessing these associations, certain relationships (especially when assessing the two smallest states in each domain) should be approached with caution.

The required number of participants will vary based on the number of latent states in each domain, and the number of time points assessed, however, consideration of this is paramount when preparing an ALTA analysis. Considering the data generated in this study, it would be useful to estimate the number of observations in each association (such as generated with the standard LTA analysis). Unfortunately in the current version, this data was not available; however, estimated sample sizes can be generated and presented through the use of a flow diagram, similar to the diagram presented in *Figure K.1*. Numerous estimates in the longitudinal assessments (*sections 9.5.3* and *9.5.4*) were '0.00' or '1.00' implying that these particular investigations could have had small sample size and as a result these estimates were interpreted with caution.

In conclusion, ALTA is a promising technique, with potential uses in many health related fields wanting to investigate longitudinal associations between two domains. However, the technique does require further development, including software, and is potentially most suitable for large sample sizes where states are prevalent to ensure reliability of parameter estimates.

The next chapter focuses on conducting sensitivity analyses on issues relating to earlier chapters of the thesis. This assesses the impact of the dichotomisation of indicators in the

hand state development, and relaxes the criteria of a minimum 5% sample size in each state, therefore permitting smaller states to be revealed. Finally, the sensitivity analyses examine the impact of removing those in the ‘least affected’ state from the analysis to explore whether other hand states were identified when the state generally without hand problems was removed.

Chapter 10: **Sensitivity analyses**

10.1 Introduction and objectives

The main objective of the sensitivity analyses was to investigate the effect of altering some of the decisions made in developing the main hand states model. In particular, whether altering the location of dichotomisation of the indicators would lead to a substantially different model. A further objective of this chapter was to explore whether permitting sample size of hand states to be less than 5%, and similarly, excluding participants with no hand pain or problem limitations, would reveal further and more detailed states that may have been missed due to the criteria specified in *Chapter 4*.

10.2 Alternative indicator dichotomisation - Methods

The population used in this analysis is identical to that used in the modelling process in *Chapter 4*. The same original 11 indicators that were highlighted by the RUG and previous literature in *section 4.5.3* were used as potential indicators for the development of the LTA model in the sensitivity analysis. In the main analysis (as detailed in *section 4.6.1*) indicators with response options 'none', 'mild', 'moderate', 'severe' and 'extreme', were dichotomised between 'mild' and 'moderate', therefore a '0' ('low') represented a response of 'none' or 'mild', whereas a '1' ('high') represented a response of 'moderate', 'severe' or 'extreme'. Through discussion with supervisors, it was agreed that the most logical alternative to this split was between 'moderate' and 'severe', therefore, in this sensitivity analysis, a '0' represents a response of 'none', 'mild' or 'moderate', while a '1' represents a response of 'severe' or 'extreme'. With regard to the indicator 'pain in both

hands', there was no adjustment, so a '0' still represented no hand pain, or hand pain in only one hand, while a '1' represented pain in both hands.

Prior to the modelling process, an investigation was carried out to display how much changes to the dichotomisation affected the frequency of individuals with a high response ('1') in each of the two datasets (for the main and sensitivity analysis). Subsequently, the modelling process that was described in detail in *Chapter 4* was undertaken.

10.3 Alternative indicator dichotomisation - Results

For the 5,617 individuals in both the main analysis and the sensitivity analysis, *Table 10.3.1* presents the proportion of participants with a high response for each of the indicators collected, based on the two dichotomisation approaches, at baseline. As mentioned in *10.2*, there were no changes to the 'pain in both hands' indicator response, which is reflected in the similar proportions in each of the databases (23%). For the other pain indicators, the mean proportion of individuals with a high response dropped from 14.9% to 4.0%. The mean proportion reporting a high response on functional indicators dropped from 13.0% to 4.6%, while the proportion with 'morning hand stiffness' dropped from 12.7% to 3.2%. These differences illustrate the impact of using a cut-point indicating a higher level of severity on the prevalence of pain and function limitations.

Table 10.3.1: Proportion of individuals with a high ('1') response for each item in the main and sensitivity analysis.

Indicator	Main analysis (n= 5,617)	Sensitivity analysis (n= 5,617)
Pain in both hands	1,289 (23.0%)[$]	1,289 (23.0%)[$]
Pain when turning objects	874 (15.7%)*	253 (4.5%)[#]
Pain when squeezing objects	970 (17.4%)*	306 (5.5%)[#]
Pain at rest	577 (10.3%)*	93 (1.7%)[#]
Pain when gripping objects	899 (16.1%)*	234 (4.2%)[#]
Difficulty opening a new jar	999 (17.8%)*	441 (7.9%)[#]
Difficulty carrying a full pot	917 (16.4%)*	372 (6.6%)[#]
Difficulty wringing out a dishcloth	855 (15.2%)*	308 (5.5%)[#]
Difficulty doing-up buttons	454 (8.1%)*	101 (1.8%)[#]
Difficulty turning taps on	420 (7.5%)*	79 (1.4%)[#]
Morning hand stiffness	715 (12.7%)*	181 (3.2%)[#]

Footnote: n= number of observations; [$]= Indicator cut-off '0'= 'left only' or 'right only', '1'= 'left and right'; *= Indicator cut-off '0'= 'none', 'mild', '1'= 'moderate', 'severe', 'extreme';
[#]= Indicator cut-off '0'= 'none', 'mild', 'moderate', '1'= 'severe', 'extreme'.

The model process for the development of hand states have been outlined in *Table 10.3.2* to *Table 10.3.5*, and is summarised in *Table 10.3.6*. The final model for the sensitivity analysis is presented in *Table 10.3.7*. The removal of only one indicator ('pain in both hands') resulted in an optimal model in the sensitivity analysis database; there was no further clear indicator to be removed from the remaining ten as the indicators with the next lowest BIC's 'difficulty opening a new jar' and 'difficulty carrying a full pot' were still deemed useful to the interpretation of the states (similar to retaining 'difficulty opening a new jar' in *section 4.7.2.2*).

Table 10.3.2: Model fit parameters with 11 indicator variables for the first removal stage.

Number of latent states	BIC	Entropy	Smallest sample size
2	54717	0.975	9.9%
3	**52052**	**0.884**	**6.8%**
4	50806	0.881	3.1%
5	50593	0.889	2.8%
6[#]	50655	0.891	1.0%

Footnote: BIC= Bayesian Information Criteria; #: Likelihood did not converge after 10,000 random starting values; Bold line represents optimum model.

Table 10.3.3: Model goodness of fit when removing one indicator in turn from model with 11 indicators and three states.

Removing the indicator:	BIC	Entropy
Pain in both hands	**37155**	**0.951**
Pain when turning objects	48611	0.876
Pain when squeezing objects	48560	0.879
Pain at rest	49821	0.881
Pain when gripping objects	48739	0.879
Difficulty opening a new jar	47424	0.880
Difficulty carrying a full pot	47579	0.880
Difficulty wringing out a dishcloth	48400	0.881
Difficulty doing-up buttons	49785	0.879
Difficulty turning taps on	50464	0.880
Morning hand stiffness	48855	0.879

Footnote: BIC= Bayesian Information Criteria; Bold line represent optimum model.

Table 10.3.4: Model fit parameters with 10 indicator variables for the second removal stage.

Number of latent states	BIC	Entropy	Smallest sample size
2	39151	0.976	8.8%
3	**37155**	**0.951**	**3.6%**
4	36868	0.948	2.5%
5	36850	0.947	1.2%
6	37090	0.930	0.8%

Footnote: BIC= Bayesian Information Criteria; Bold line represents optimum model.

Table 10.3.5: Model goodness of fit when removing one indicator in turn from model with 10 indicators and three states.

Removing the indicator:	BIC	Entropy
Pain when turning objects	33913	0.946
Pain when squeezing objects	33883	0.948
Pain at rest	34961	0.949
Pain when gripping objects	34087	0.945
Difficulty opening a new jar	32712	0.944
Difficulty carrying a full pot	33019	0.943
Difficulty wringing out a dishcloth	33839	0.945
Difficulty doing-up buttons	35036	0.951
Difficulty turning taps on	35744	0.948
Morning hand stiffness	33939	0.950

Footnote: BIC= Bayesian Information Criteria.

When comparing the LTA model selected in the sensitivity analysis with that of the main analysis, there were two main differences. Firstly, the indicators included were not identical between the two models; the sensitivity analysis contains a total of ten. Therefore, the sensitivity analysis included the additional indicators ‘pain at rest’ and ‘morning hand stiffness’. It may be of interest to recall that these two items were the last ones to be removed from the total numbers of items in *section 4.7.2.2*.

Table 10.3.6: Modelling process for optimal LTA model in sensitivity analysis database.

Removal stage	Number of indicators	Indicator removed	Number of phenotypes	BIC after removal	Entropy after removal	Smallest state sample size
1	11	-	3	52052	0.884	6.8%
2	**10**	**Pain in both hands**	**3**	**37155**	**0.951**	**3.6%**
3	9	Difficulty opening a new jar	3	32712	0.944	4.9%

Footnote: BIC= Bayesian Information Criteria; Bold line represents optimum model.

Secondly, the optimum number of states was only three, in contrast to the main analysis which had five states as optimum. When considering the item-response probabilities, the states had some similarities with those created in the main analysis (*Table 10.3.7*). The first state had very low probabilities for all indicators, which was analogous to the 'least affected' state in the main analysis. The second state had two functional indicators ('difficulty opening a new jar' and 'difficulty carrying a full pot') that showed a high probability of positive responses. These two indicators and the pattern of the other indicators were comparable to the state previously labelled 'poor gross function'.

The third state encompassed less than 6% of the population at each time point and all indicators had higher probabilities compared to the other states (although 'pain at rest', 'morning hand stiffness', 'difficulty doing-up buttons' and 'difficulty turning taps on' had probabilities less than 0.5). The characteristics of this state were similar to those of the 'high pain and poor gross function' state of the main analysis.

Table 10.3.7: The optimal LTA model for the sensitivity analysis database.

(n= 5,617)	State		
	1	2	3
Potential labels:	Least affected	Poor gross function	High pain & poor gross function
Latent state proportions			
Baseline	0.893	0.071	0.036
3 years	0.853	0.095	0.053
6 years	0.857	0.100	0.043
Item-response probabilities			
Pain when turning objects	0.000	0.208	0.814
Pain when squeezing objects	0.001	0.255	0.958
Pain at rest	0.001	0.085	0.253
Pain when gripping objects	0.000	0.197	0.747
Difficulty opening new jar	0.003	0.549	0.939
Difficulty carrying full pot	0.003	0.414	0.879
Difficulty wringing out a dishcloth	0.001	0.299	0.859
Difficulty doing-up buttons	0.000	0.059	0.369
Difficulty turning taps on	0.000	0.041	0.305
Morning hand stiffness	0.004	0.150	0.470
Latent transition probabilities			
Baseline to 3 years	Least affected	Poor gross function	High pain & poor gross function
Least affected	**0.934**	0.051	0.015
Poor gross function	0.211	**0.595**	0.194
High pain & poor gross function	0.035	0.297	**0.668**
3 years to 6 years	Least affected	Poor gross function	High pain & poor gross function
Least affected	**0.949**	0.046	0.005
Poor gross function	0.389	**0.501**	0.109
High pain & poor gross function	0.209	0.298	**0.492**

Footnote: Bold entries represent stability between the time points.

When considering the transitional probabilities in the sensitivity analysis, there were high stability probabilities for the 'least affected' state at each time point; this finding was similar to the transition probabilities in the main analysis. In addition to this, the stability of the 'poor gross function' and 'high pain and poor gross function' states were not as high as the stability in the 'least affected' state, similar to the main analysis. The patterns between baseline to 3 years, and 3 years to 6 years were reasonably similar, with the largest discrepancy in a higher probability of individuals transitioning from 'high pain and poor gross function' to 'least affected' at 6 years, than at 3 years (τ= 0.209 and τ= 0.035 respectively).

10.4 Additional hand state - Methods

One potential limitation of restricting the state population to at least 5% (as defined in *section 4.6.4.3*) was that less prevalent states of hand pain/ problems which may have represented important clinical groups could be missed. In an attempt to explore whether relevant states were missed due to this criterion, a sensitivity analysis permitting state sample size to be smaller than 5% was performed using same indicator terms included in the base model using LTA. Therefore, compared with the five state base model displayed in *Table 4.8.1*, a model with six hand states was explored in terms of the hand pain and function profiles.

10.5 Additional hand state - Results

The LTA model was developed for six hand states using the same eight indicators highlighted in *section 4.7.2.2,* and is displayed in *Table 10.5.1*. Investigating the item-response probabilities, the first five hand states had a similar profile to the five states developed in *Chapter 4*, and were given similar labels. The new sixth hand state is presented on the right of *Table 10.5.1* and is emphasised in bold. This state is made up of 1.8% of the participants at baseline, and the participants in this state were mainly affected by the poor gross function related indicators (all ρ>0.82), whilst participants also reported 'pain when squeezing objects' and 'difficulty turning taps on' and 'difficulty doing-up buttons' (0.37>ρ<0.42). However, these three indicators were in (or close to) the region of 0.4 to 0.6, which implied they were not well defined in that state (Collins and Lanza, 2010). For example, an indicator of ρ= 0.5 would indicate that half of the participants in that state would be expected to score high on indicators, while the others would not.

The additional hand state appeared to be made up of individuals previously assigned to the 'poor gross function' and 'high pain and poor gross function' states, and therefore the new

state could be regarded as a subset of the previously determined two states. It was decided with the supervisory team that in this scenario, this additional hand state was not well defined and the original five hand state model was deemed optimum.

Table 10.5.1: State proportion and item-response probabilities of six hand state model.

	State					
	1	2	3	4	5	**6**
Latent state proportions at baseline	0.759	0.040	0.049	0.064	0.069	**0.018**
Potential labels:	LA	HP	PGF	HPPGF	SA	**PGF + pain squeezing**
Pain when turning objects	0.000	0.639	0.105	0.880	0.983	**0.198**
Pain when squeezing objects	0.002	0.854	0.083	0.930	0.999	**0.410**
Pain when gripping objects	0.005	0.711	0.148	0.848	0.958	**0.169**
Difficulty opening a new jar	0.004	0.178	0.559	0.848	0.999	**0.990**
Difficulty carrying a full pot	0.004	0.063	0.492	0.787	0.979	**0.920**
Difficulty wringing out a dishcloth	0.002	0.150	0.312	0.735	0.987	**0.823**
Difficulty doing-up buttons	0.001	0.025	0.041	0.171	0.858	**0.377**
Difficulty turning taps on	0.000	0.018	0.040	0.068	0.878	**0.423**

Footnote: LA= 'least affected'; HP= 'high pain'; PGF= 'poor gross function'; HPPGF= 'high pain and poor gross function'; SA= 'severely affected'; Bold column represents newly identified latent state.

10.6 Excluding the 'least affected' state - Methods

77% of the participants (δ= 0.769, *Table 4.8.1*) were estimated to be in the 'least affected' hand state at baseline. A potential implication of this, similar to the previous analyses, was that other smaller states may not be discovered as the 'least affected' state was estimated to contain a high proportion of participants. Therefore, an additional analysis was performed excluding the individuals that were classified in 'least affected' at baseline, and then the

hand states remodelled using LTA, again with the same indicators contained in the base model in *section 4.8*. The optimum number of states was explored to identify the most favourable model, and how this compared to the model defined in the main analysis (*Table 4.8.1*).

10.7 Excluding the 'least affected' state - Results

A summary of the model criterion for various numbers of hand states is displayed in *Table 10.7.1*. The models with four and five hand states appeared to be the most favourable from the summary of results. The model with four hand states had a reasonable balance of lowest BIC, high entropy and substantial sample size, whereas five hand states had a slightly improved BIC, but lower entropy and a smaller sample size around the 5% boundary.

Table 10.7.1: Summary of model criterion for different number of hand states with no 'least affected' state.

Number of states	BIC	Entropy	Smallest sample size
2	27,869	0.907	42.7%
3	25,847	0.868	24.9%
4	25,338	0.866	17.5%
5	25,146	0.858	5.9%
6	25,274	0.868	5.7%

Footnote: BIC= Bayesian Information Criteria.

The state proportions and item-response probabilities of the model with four hand states is displayed in *Table 10.7.2*. The profiles of the four hand states were analogous to the problematic hand states defined in the base model displayed in *section 4.8*, and mirrored the 'high pain', 'poor gross function', 'high pain and poor gross function' and 'severely affected' states.

Table 10.7.2: Latent state proportion and item-response probabilities of four hand state model with no 'least affected' state.

	State			
	1	2	3	4
Latent state proportion at baseline	0.175	0.243	0.270	0.313
Potential labels	HP	PGF	HPPGF	SA
Item-response probabilities				
Pain when turning objects	0.674	0.154	0.888	0.958
Pain when squeezing objects	0.858	0.188	0.943	0.989
Pain when gripping objects	0.745	0.166	0.850	0.938
Difficulty opening a new jar	0.230	0.747	0.845	0.999
Difficulty carrying a full pot	0.047	0.690	0.813	0.979
Difficulty wringing out a dishcloth	0.151	0.508	0.756	0.986
Difficulty doing-up buttons	0.021	0.127	0.178	0.857
Difficulty turning taps on	0.018	0.138	0.067	0.880

Footnote: HP= 'high pain'; PGF= 'poor gross function'; HPPGF= 'high pain and poor gross function'; SA= 'severely affected'.

Table 10.7.3 displays the model with five hand states excluding participants in 'least affected' at baseline. The profiles of the five states were similar to those defined in the previous sensitivity analysis displayed in *Table 10.5.1* (*section 10.5*). A similar problem to the model defined in *10.5* was the overlap between the additional state (state five in *Table 10.7.3*) and two existing states, 'poor gross function' and 'high pain and poor gross function', and the weak distinction of 'pain when squeezing objects', 'difficulty doing-up buttons' and 'difficulty turning taps on' ($0.45<\rho>0.56$) in the fifth state. As a consequence of this relatively weak distinction, the model with four hand states was the most favourable when not including 'least affected'. Therefore, this provides some reassurance that distinct and clinically-meaningful hand states were not missed as a result of a large proportion of individuals being classified in the 'least affected' state at baseline in the main analysis.

Table 10.7.3: State proportion and item-response probabilities of four hand state model with no 'least affected' state.

	State				
	1	2	3	4	5
Latent state proportions at baseline	0.177	0.188	0.278	0.299	0.059
Potential labels	HP	PGF	HPPGF	SA	?
Item-response probabilities					
Pain when turning objects	0.666	0.144	0.879	0.985	0.186
Pain when squeezing objects	0.859	0.117	0.937	0.999	0.457
Pain when gripping objects	0.735	0.160	0.845	0.959	0.200
Difficulty opening a new jar	0.196	0.704	0.855	0.999	1.000
Difficulty carrying a full pot	0.083	0.642	0.793	0.979	0.897
Difficulty wringing out a dishcloth	0.162	0.434	0.750	0.987	0.822
Difficulty doing-up buttons	0.027	0.067	0.177	0.859	0.456
Difficulty turning taps on	0.018	0.058	0.070	0.882	0.559

Footnote: HP= 'high pain'; PGF= 'poor gross function'; HPPGF= 'high pain and poor gross function'; SA= 'severely affected'.

10.8 Summary

The first sensitivity analysis in this chapter showed that changing the cut-point for the dichotomisation of the indicators yielded a different model, resulting in three rather than five hand phenotypes. This indicated that the model produced in the main analysis was not entirely robust to changes in the cut-point used for dichotomising items. However, using a minimum 5% state proportion cut-off, and including individuals with no/ little hand pain or function symptoms did not miss potentially relevant states.

The differences found in the first sensitivity analysis were likely to be explained by the fact that the proportions of individuals with a positive response decreased considerably in the sensitivity analysis once individuals had to report severe or extreme pain or functional issues to be considered as having hand problems. Therefore, the information available to

distinguish between different hand characteristics in the modelling process was limited in comparison to the main analysis. This also provided an explanation why more indicators were included in the model (as there was less information for the modelling process to decide which indicators were unnecessary), and also the reason why a lower number of states were optimum, as the majority of individuals did not report any issues based on the adjusted dichotomisation (646 reported at least one of the eight baseline complaints (12%) in the sensitivity analysis, compared to 1,397 (25%) in the main analysis using the original dichotomisation).

Investigating the impact of including states with sample sizes less than 5% provided reassurance that the five hand states developed in *Chapter 4*, and used throughout this thesis, was the optimal model. The next additional state appeared to be a subset of the 'high pain and poor gross function' state and contained indicators with item-response probabilities in the region of 0.4 to 0.6, which would suggest a lack of homogeneity in this state. In addition to this, exploring the effect of removing the participants in the 'least affected' state at baseline did not identify any hidden hand states.

A limitation of the two latter sensitivity analyses was that the modelling process was not completed from the beginning, as was done for modifying the indicator cut-off point (*section 10.2*). It was possible that different indicators (or a different number) could have been selected for the final model. However, for the analysis of permitting states less than 5%, it is unlikely that different indicators would have been selected as the removal of indicators was mainly based on the BIC and entropy (*section 4.7.2.2*). Similarly, for the sensitivity analysis removing those in the 'least affected' hand state at baseline, the prior modelling development was required in order to identify those who were 'least affected' and then subsequently exclude them.

A further potential additional sensitivity analysis could have been to remove individuals who reported no hand pain/ problems at all time points. However, in the main analysis the 'least affected' state also included 918 individuals at baseline (twice as many individuals as in any other state at baseline) who reported hand pain/ problems in the HS and RPS (897 at 3 years, 665 at 6 years), but then reported 'none' or 'mild' symptoms to the indicators in the base model (*section 5.6*). Due to the substantial number of individuals that would still be included the LTA process, it is unlikely that excluding individuals reporting no hand pain/ problems at each time point would change the defined states.

The next chapter, the discussion, summarises the main findings produced during this project, as well as relating these findings to other literature. In addition, the strengths and limitations of the applied research, and methodology used is addressed, before providing potential research objectives for future research, and the implications of the research completed.

Chapter 11: **Discussion**

In this PhD project I have developed phenotypes of hand pain and problems in older people and explored their transitions over time using a Latent Transition Analysis approach to increase the knowledge on the long-term course of hand pain and function, and investigate methodology rarely utilised in MSK research. This chapter reviews the key findings from the thesis and highlights the strengths and limitations of this work, firstly addressing the clinical findings, and secondly, the statistical techniques applied. Potential future research and the implications of this work are subsequently considered.

11.1 Summary of main findings

11.1.1 Objective 1: To identify distinct states of hand pain and function

Latent states of hand pain and hand function were developed (***Chapter 4***) based on key items that were highlighted as being important by older individuals with hand problems and as indicated by previous literature. These states classified individuals into profiles that represented distinct patterns of hand pain and function, which were labelled: 'least affected', 'high pain', 'poor gross function', 'high pain and poor gross function' and 'severely affected'. These states were well defined, in that they each reflected a different profile of hand symptoms and no two states were overly similar. Sensitivity analyses altering the point of dichotomisation of the indicators so that individuals had to have more severe problems to be regarded as having a positive response to an indicator, exploring additional hand states, and removing those in the 'least affected' state at baseline, did not suggest any distinct improvement, indicating that the model was robust to these changes (Chapter 11). The performance of this model was evaluated (***Chapter 5***) by comparing the

distribution of observed indicator responses to the estimated item-response probabilities, and assessing the average posterior probabilities to ensure individuals had been clearly classified into their optimum hand state. The exploration of these measures revealed that the states were logical (based on the indicator responses), and had high average posterior probabilities for state membership. Finally, it was reasonable to restrict the item-response probabilities over time to allow state definitions to remain constant at each time point, therefore permitting a more straightforward interpretation of transitions between the states.

11.1.2 Objective 2: To investigate the transitions of individuals between the hand states over time

The longitudinal pattern of hand pain/ problems over the 6 years follow-up period, through the exploration of transition probabilities, was assessed in detail in ***Chapter 6***. In brief, the main finding was that the probability of staying in the same hand state was high at either end of the spectrum at each of the follow-up time periods ('least affected' ≥0.86, 'severely affected' ≥0.68). However, the estimated probability of transitioning from 'high pain' to 'least affected' was notably large (τ= 0.42 baseline to 3 years, τ= 0.48 3 years to 6 years), and those in hand states with function issues were less likely to improve at subsequent follow-up points compared to those with pain alone. Further investigations suggested initially females and those with widespread pain were more likely to have unfavourable transition patterns, however, including covariates when developing the latent states was not deemed to be a necessary inclusion.

11.1.3 Objective 3: To explore factors predicting future hand state membership

Multivariable analysis (***Chapter 7***) revealed those with nodes, sleep problems, chronic hand pain, bilateral hand pain, poor self-reported general health and in younger age groups

at baseline were more likely to be in an unfavourable hand state at 3 years. Baseline predictive factors of 6 year hand state membership were similar, although females were also more likely to be 'severely affected' compared to males. Similar factors (including nodes, female gender, chronic duration, bilateral hand pain) appeared to predict more unfavourable trajectories ('deterioration' or 'stability of problematic conditions') in those who were in the intermediate and worst hand states at baseline.

11.1.4 Objective 4: To investigate the associations of hand states with primary care consultation and medication prescription

Exploring the linked health care data within the NorStOP study (***Chapter 8***) revealed that many individuals with hand pain/ problems did not appear to utilise health care over the 6 year period. Only 13% of those who consented to medical record review consulted for a hand-related problem, with 22% of those in the worst hand state ('severely affected') at baseline consulting over the following 6 years. Baseline hand state membership predicted consulting for a hand-related condition better than basic demographic information. However, hand-specific factors such as nodes and frequent medication use for hand symptoms were most highly associated with consultation. A third (33%) of consulters received a prescription within 14 days of a hand-related consultation, with the majority receiving non-opioid analgesia (21% of those who consulted).

11.1.5 Objective 5: To explore the longitudinal relationship between hand states and mental health states

Mental health phenotypes were developed using a similar approach to that used for the hand states, and subsequently the longitudinal association between the mental health and hand states was assessed with the novel technique of ALTA (***Chapter 9***). The four identified mental health states were labelled: 'no anxiety/ depression', 'mild depression',

'high anxiety', 'anxiety and depression'. Those in the healthier hand states were more likely to be (and remain) in the healthier mental health states over the 6 years follow-up. Conversely, those in the more severe hand states were also more likely to be (and remain) in the more severe mental health states. A potential relationship was discovered between pain and depression, and with function and anxiety, however these relationships require further exploration to determine if associations with hand states are indeed different for anxiety versus depressive symptoms.

11.2 Comparison to previous literature

11.2.1 Prevalence of hand pain/ problems

Of the 5,617 participants analysed in this project, 2,197 (39%) reported having at least one hand pain or function problem at baseline. However, only 1,279 of these were classified in a problematic hand state at baseline, which led to an overall point prevalence for hand pain/ problems of 23% for the sample. This prevalence estimate was high compared to other published estimates of the point prevalence of symptomatic hand OA, which tend to be less than 10% for populations aged around 50 years and older (Aihie Sayer et al., 2003; Zhang et al., 2003; Dillon et al., 2007; Grotle et al., 2008a; Massengale et al., 2012; Yeşil et al., 2014). While it is expected that the majority of participants in this study reporting hand pain/ problems had some degree of hand OA based on their age and hand symptoms, it is likely that a proportion of these had other types of hand conditions (discussed later). In addition to this, a proportion of the participants may not meet the stricter definitions of hand OA used in these other studies, which may partly explain the higher estimate for the prevalence of hand pain/ problems in this study.

11.2.2 Indicators of hand pain and function symptoms

The identification of important indicators related to hand pain and problems based on patient opinions and previous literature indicated that most of the items in the AUSCAN scale were of interest and were therefore used as potential indicators. An exception to this was 'pain at rest', which was determined to be unnecessary (based on statistical criteria and state interpretation) and was removed as a potential indicator during the model development. The removal of this item in a previous Rasch analysis has been found to improve the AUSCAN pain scale, with the conclusion that 'pain at rest' reflected a separate pain construct to pain when doing an activity (which relate to the other four pain items) (Haugen et al., 2011b). In addition to this, the same study concluded that dividing the function items into two constructs, 'grip strength' and 'high precision', improved the dimensionality of the AUSCAN function subscale (Haugen et al., 2011b). These two findings were highly analogous to the results of the model development in this study, where 'pain at rest' was not included as an indicator, and two main function profiles were evident. This was apparent based on the indicators labelled 'poor gross function' (similar to 'grip strength' in Haugen et al., 2011b), and the indicators of poor fine motor skills (similar to 'high precision' in Haugen et al., 2011b), which were the main differences between the 'high pain and poor gross function' and 'severely affected' hand states. However, Haugen and co-authors deemed the item 'difficulty turning taps on' to be attributed to 'grip strength' ('poor gross function'), while in this population it was clustered with indicators of poor motor skill. Due to the various designs of taps in different populations (already discussed in *section 7.6.1*), it is conceivable that this item could span both of these function constructs. From this study, and that performed by Haugen et al., it appears that the AUSCAN questionnaire could benefit from the function scale being regarded as two functional constructs (rather than the current one dimension) when

computing subscale scores, and also potentially from the removal of 'pain at rest', which has been shown to provide little additional information.

11.2.3 Symptom course over time

The transition probabilities found that participants with pain but no functional limitations were more likely to see an improvement in their symptoms compared to those with function issues. Other longitudinal studies addressing changes in symptomatic pain and function in older individuals with hand pain/ problems are sparse. However, the findings in this thesis have similarities to trajectory work using LCGA in other OA locations, such as the knee and hip (Verkleij et al., 2012; Collins et al., 2014; Nicholls et al., 2014). These studies found groups of individuals with pain symptoms did indicate signs of improvement in their OA condition over the study period.

In a study of 289 participants with confirmed radiographic hand OA, individuals were more likely to deteriorate in terms of their function scores of the AUSCAN (Bijsterbosch et al., 2011). 50% of patients reported more functional difficulties after 6 years of follow-up, and 40% of patients reporting more pain. However, the proportion of individuals reporting an improvement in their pain or in function scores was identical (26%).

A study of individuals consulting for hand and wrist problems in primary care found that 42% of individuals reported complete recovery at 1 year after first consulting their GP (Spies-Dorgelo et al., 2008). In this study, 23% were classified in the 'least affected' at 3 years having been not been in 'least affected' at baseline (296/ 1,279). However, Spies-Dorgelo and colleagues included adults aged 18 and older, with a broad range of hand/ wrist problems (only 17% with a diagnosis of osteoarthritis), and at an age more susceptible to recovery.

11.2.4 Predictive factors of progression of hand pain/ problems

While previous state membership was in most cases the strongest predictive factor of current state, the analysis in this thesis highlighted some factors that were predictive of future hand state membership. Many of these have been established in previous literature as factors associated with the onset of hand pain/ problems, but knowledge of whether these factors are predictive of the progression of hand conditions is scarce. The majority of discussion regarding the predictive factors has been presented at the end of *Chapter 7* (*section 7.6*), however some of the main results are summarised below.

The predictive factors that were found to be associated with the progression of hand pain/ problems were generally similar to previous literature. For example, chronic hand pain duration was found in this project to predict a more unfavourable outcome at 3 and 6 years. Also, in a sample of adults consulting for hand pain/ problems in a Dutch general practice, chronic hand symptom duration was also associated with a poor outcome at 3 and 12 months (Spies-Dorgelo et al., 2008). This evidence suggests that long duration of symptoms is associated with poorer prognosis of hand pain/ problems.

Females aged 18 years and older have been found to have a poorer outcome than males at 3 and 12 months after consulting for hand problems (Spies-Dorgelo et al., 2008). However, the relationship between gender and the progression of the condition does not appear as strong as the relationship between gender and the onset of hand pain/ problems (van Saase et al., 1989; Zhang et al., 2002; Zhang et al., 2003; Visser et al., 2014). Therefore, from findings in previous work and predominately the results in this study, while female gender has been strongly associated with the onset of hand pain/ problems, the association with progression is not as strong.

While nodes are commonly seen in individuals with hand OA (Alexander, 1999; Jones et al., 2001; Thaper et al., 2005; Dillon et al., 2007; Bijsterbosch et al., 2011; Kwok, 2013; Ghosh et al., 2014), little research has explored the association between the presence of hand nodes in individuals with hand pain/ problems and long-term outcomes. In the 6 year prospective study by Bijsterbosch et al., nodes were present in 71% of the population with radiographic hand OA (Bijsterbosch et al., 2011). The presence of nodes was mildly associated with poor long term pain outcome and strongly associated with radiographic progression, but not with functional outcome (Bijsterbosch et al., 2011). In this thesis, 55% of the most severe state reported the presence of nodes at baseline (*Appendix F*), and it was highly predictive of an unfavourable hand state at follow-up. Therefore, this highlights to clinicians and researchers that individuals with nodes are more likely to have an unfavourable long-term outcome.

Other studies looking at cross-sectional or progression of hand OA/ pain/ problems have tended not to report on the presence of bilateral hand pain. The development of (radiographic) hand OA is often bilateral, therefore these conditions frequently develop across both hands, than in one hand alone (Marshall et al., 2009). In addition to this, the diagnosis of symptomatic hand OA tends to use information on both hands (Altman et al., 1990). Consequently, using bilateral hand pain symptoms as a predictive factor may not be the most suitable measure of hand severity. Similar research has tended to use the number of OA affected hand joints as a predictor of outcome, where a higher number of affected joints predicted a poorer outcome (Bijsterbosch et al., 2011), which is a similar finding in this research using bilateral hand pain as a predictive factor. A similar factor was included in the analysis (‘pain in two or more hand joints’), which was non-significant in all multivariable analyses. The focus of this factor is distinctly different to the ‘number of OA

joints' which is used in other literature. Consequently bilateral hand pain could be regarded as a more relevant proxy for assessing the extent of hand symptoms.

While the older age groups were more likely to be in the more severe hand states at baseline, this relationship was not consistent when exploring the progression of hand pain/ problems. Several explanations for this finding have been previously discussed (*section 7.6*), but perhaps most notable is that this may reflect that older people adapt their everyday lifestyle to cope with their hand conditions (Myers, 2008; Bukhave and Huniche, 2014). While logical, the potential for adaptation to explain observed improvements of hand conditions in older adults requires further investigation for firm conclusions to be made.

Sleep problems were found to be a consistently significant predictive factor of long-term severe hand state membership in the population analysed in this thesis, however this predictive factor has rarely been considered in the literature. Poor sleep quality associated with presence of hand OA was highlighted in a cross-sectional Norwegian study (Grotle et al., 2008a), but it was not reported in any prospective cohort studies in the literature looking at onset or progression of hand problems. Future research should seek to include a more detailed assessment of sleep problems to evaluate the extent to which sleep disturbance can help to identify individuals at risk of developing a more severe hand condition.

Overall, the results from this thesis confirm and support those of the previous literature and have revealed a consistency of findings. The factors included in the analysis have previously been highlighted, mainly for their cross-sectional associations with severity of hand pain/ problems, and the longitudinal associations found in this study provides additional validation to the hand phenotypes identified. Ideally the results found in this

study would be compared with other research exploring factors of progression of hand pain/ problems over time; however, as highlighted by a recent systematic review, very few studies have investigated predictive factors of progression in hand pain/ problems (Nicholls et al., 2012). Therefore, it can be considered that the findings contained in this PhD present new evidence to the field of hand pain/ problems in older people in the general population.

11.2.5 Primary health care for hand pain/ problems

The proportion of individuals consulting for hand problems in the NorStOP population was low (13%), and of those in the more severe states at baseline, only a fifth of individuals consulted over the 6 years follow-up period. Of those who did consult, a third received an analgesia prescription. This prescription rate is very similar to that found in a Dutch consulting population, where 36% of patients were prescribed medication on their first consultation for hand problems (Spies-Dorgelo et al., 2008). In an 8 year follow-up of British individuals with OA in various locations, NSAID use was seen to decrease over the follow-up period (from 58% to 29%), with other analgesia use increasing (from 30% to 54%) in those with confirmed hand OA (Dieppe et al., 2000), potentially due to guidance and increasing concerns over side-effects from oral NSAIDs. The sample size in this thesis (for the health care analysis) is limited by the small number of individuals that consulted with hand problems, thus, it was difficult to emphasise any trends investigating specific analgesia prescriptions. In terms of the proportion of individuals receiving a prescription at 6 years compared to baseline, there was a suggestion of a slight increase of prescriptions in the states containing those with hand pain and dysfunction.

11.2.6 Longitudinal associations between hand pain/ problems and mental health

Previous research has investigated the relationship between musculoskeletal conditions and depression/ anxiety (Arola et al., 2010) and demonstrated that baseline bodily pain (any location) was a significant predictor of developing depression (OR= 2.47 (1.96, 3.11)) and anxiety (OR= 2.02 (1.60, 2.55)) at 3 years follow-up (Arola et al., 2010). Similarly, the authors found that baseline depression (OR= 2.42 (1.24, 4.69)) and anxiety (OR= 2.30 (1.67, 3.17)) were significant predictors of bodily pain (any location) at 3 years follow-up (Arola et al, 2010). In addition to this, Kroenke et al., found that change in pain was a strong predictor of depression severity at follow-up (p<0.0001), while change in depression severity was also a strong predictor of pain severity at follow-up (p<0.0001), at 3, 6 and 12 months (Kroenke et al., 2011).

A strong association between the most severe hand and mental health states, and the least severe hand and mental health states was observed in the findings presented in this thesis. These findings are comparable to the overall evidence for associations between MSK and mental health (individuals are often affected by both conditions simultaneously, Demyttenaere et al., 2007; Mallen et al., 2007; Kroenke et al., 2011).

The analysis presented in this thesis highlighted relationships between pain and depression, and between poor function and anxiety, both cross-sectionally and longitudinally. Due to the limited research on these specific areas in hand studies, comparison to other literature is difficult, however associations between pain and depression have consistently been found for MSK conditions in a more general context (Bair et al., 2003; Arola et al., 2010; Kroenke et al., 2011). In addition, anxiety has been found to be predictive of knee function disability over an 18 month period (Mallen et al., 2007). As these specific relationships

exist in other MSK conditions, further exploration with respect to hand conditions is warranted.

11.3 Strengths and limitations

This project was a large study of hand pain/ problems in the general population of older adults, with long term follow-up using LTA. A substantial strength of this project was this was the first time a statistical technique has been employed to discover underlying groups of individuals in the population in reference to both their hand pain and function, whilst exploring their respective patterns over time. Previous research has tended to explore mean change in pain and function separately, whereas the latent variable methodology used here has uncovered sub-groups in the population that would not be obvious otherwise. For example, analysis of pain or function separately would not have revealed the 'high pain and poor gross function' state. This state reflected a combination of two complaints and potentially contained individuals who were at a different stage of hand problems or represented a different subset of hand problems compared to individuals who only reported one of these dimensions.

11.3.1 Study population

Sampling from general practice in the UK allows identification of a sample representative of the general population as 98% of the UK population is registered with a general practitioner (Bowling, 1997). Inclusion criteria in this population was based on being aged 50 years and older, and registered at a general practices. This provided a broad opportunity to capture hand conditions with widely ranging levels of severity, as opposed to only including individuals who consult for example (Spies-Dorgelo et al., 2008).

The participants used for this study were registered at one of eight general practices in North Staffordshire in the UK, and as a result this population may not be representative of the UK population as a whole, or other international populations. Nevertheless, the general practices used in the NorStOP cohort were sampled to represent areas with different deprivation levels, and contained in rural and non-rural areas. In addition to this, the consultation prevalence of MSK conditions in North Staffordshire has been shown to be similar to that in the UK (Jordan et al., 2007).

11.3.2 Loss to follow-up

A total of 5,751 individuals responded at all three time points (31% of baseline responders), and therefore the majority of the baseline responders (n= 12,746) were not included in the main analysis. Baseline comparisons revealed those who were not followed-up over the 6 years were more likely to be older, retired, or lived alone, but there was little difference by gender (*Table 4.7.1, section 4.7.1*). A potential consequence of the absence of older, 'single' participants in the study sample could be an underestimation of the prevalence of the more severe hand pain/ problems states as older individuals are generally more at risk of hand pain/ problems and not having others to assist may accentuate their problems. Therefore, the lower proportion of those more susceptible to hand pain/ problems in this sample could result in more 'optimistic' transition patterns over the follow-up period, and dilute the magnitude of association between predictive factors and future state membership (those more severe were not included in the modelling processes). An additional analysis including individuals that responded at 3 years but not at 6 years (using an additional 4,088 individuals, 9,705 in total), found few differences developed between the profiles of the states or the proportions of individuals in each hand state (*section 4.7.2.1*). Therefore this provides some reassurance that differences between

the study sample and the source population did not greatly bias the results of the LTA modelling process.

11.3.3 Base model development

The development of the hand state model in this project commenced with valuable input and opinions from individuals with hand pain/ problems, explained in *Chapter 4*. Item selection using existing evidence and expert input (patients and clinicians) were arguably more likely to provide more generalisable results compared to item selection purely based upon statistical findings (which would likely be influenced by missing values and sampling variation). Using a purely statistical approach to narrow down the number of indicators from the initial 40 may have led to the exclusion of indicators that were viewed as important by individuals with hand pain/ problems.

The questions and factors included in this project were restricted to the information collected in the NorStOP study. While the NorStOP study included a large amount of information, some items were not collected at all time points. For example, five items were removed from the model development stage in *Chapter 4* because these were not collected at 6 years follow-up (which were 'hand pain in two or more joints', 'write easily with a pen', 'burning sensation', 'hand stiffness' and 'hand numbness'). While these were not available at 6 years, the 3 years pilot analysis revealed that these items were not required for the final model which suggested these items would likely have been irrelevant in the final 6 year model had they been collected. The RUG did highlight some items that were not directly assessed in the NorStOP database, such as previous occupation containing heavy physical work load, and the presence of hand OA in parents (occupational and familial factors); however this information would be more useful to use as a predictive factor of future state membership (next section, *section 11.3.4*).

11.3.4 Predictive factors of state membership

The NorStOP study had a wide range of potential factors available to predict long term hand state membership but several factors that might be associated with future hand states were not included. For example, education was not included as a potential factor. The main reason for this was previous work on the NorStOP population revealed a higher than expected proportion continuing to further education (Lacey et al., 2012; McBeth et al., 2014), so the accuracy of this question was doubtful. Further, ethnicity of the participant was not included due to the small proportion of ethnic minorities recruited in the study (97% white).

Further factors that could have been collected to predict hand problems (based on some of the previous literature) include family history of hand conditions (mentioned previously) and more specific information on individual's weight and body shape, such as waist-to-hip ratio, waist circumference, body type (for example, athletic, slender) which have been found in some studies to be related to hand problems (Kalichman and Kobyliansky, 2007; Visser et al., 2014). Further to this, previous work has identified a relationship between post-menopausal status in women, and the development and progression of hand OA (Kalichman and Kobyliansky, 2007; Prieto-Alhambra et al., 2014; Yeşil et al., 2014). The majority of these factors are less easy to assess using self-reported data (e.g. waist-to-hip ratio and waist circumference), or incur recall bias (e.g. remembering number of years since menopause), therefore the reliability of the responses could have been questionable if they had been included in the NorStOP study.

11.3.5 Hand osteoarthritis diagnosis

A potential limitation of this work included the absence of x-rays or physical examinations, as they were not available so confirmation as to whether participants with problematic

hand conditions had radiographic or physician-diagnosed hand OA could not be made. However, previous work has revealed that in the majority of older individuals with hand problems, radiographic features of osteoarthritis can be found (van Saase et al., 1989; Jones et al., 2001; Riyazi et al., 2005; Marshall et al., 2013). For example, in one population sample of 343 people aged between 55 and 76 years with hand problems, 92% had radiographic hand OA after a 24 year follow-up (Chaisson et al., 1997). In addition to this, a study assessing a subset of the NorStOP population (CAS-HA study, Myers et al., 2007) with more detailed investigations found that radiographic hand OA was present in 78% of individuals with hand pain (Marshall et al., 2013). Therefore, it is conceivable that the majority of older individuals with problematic hand conditions would have had some radiographic change.

However, analysing radiographs for MSK conditions has been questioned in the literature due to inconsistent findings between self-reported pain and function scores with radiographs (Hannan et al., 2000; Felson and Nevitt, 2004). For example, changes in self-reported pain and function were not significantly associated with radiographic progression in a secondary care sample of patients with hand OA, although follow-up was relatively short (2 years) (Botha-Scheepers et al., 2009). With these concerns in mind, assessing pain and function symptoms is of more relevance to patients, and should be the primary focus in MSK research (Felson and Nevitt, 2004).

It is conceivable that individuals with other hand conditions aside from hand OA were included in this data set. For example, participants could have had RA, carpal tunnel syndrome, Dupuytren's contracture, trigger finger or DeQuervain's tenosynovitis. In the subset of 578 NorStOP participants analysed in the CASHA study containing older individuals reporting hand pain/ problems, other hand conditions were less frequent than OA (carpal tunnel syndrome 46%, Dupuytren's contracture 26%, DeQuervain's

tenosynovitis 23%, trigger finger 20%), while over 80% of participants were diagnosed with hand OA (Marshall, 2010; Marshall et al., 2013). In a study assessing all adults consulting with hand problems in primary care, 17% of participants were diagnosed with osteoarthritis, 16% with tenosynovitis, 12% with carpal tunnel syndrome and 8% with RA (Spies-Dorgelo et al., 2007). However, this study also included younger adults and recruited individuals consulting for hand pain/ problems, so may have therefore provided a different spectrum of hand conditions compared to the population-based sample of older adults used in this thesis.

The predictive analysis contained within Chapter 7 highlighted that nodes and bilateral hand pain were highly predictive of membership in more severe hand states. Therefore as these two aspects are common characteristics of hand OA (Alexander, 1999; Jones et al., 2001; Zhang et al., 2002; Thaper et al., 2005; Dillon et al., 2007; Bijsterbosch et al., 2011; Kwok, 2013; Ghosh et al., 2014) but not of other common hand conditions (carpal tunnel syndrome, Dupuytren's contracture, DeQuervain's tenosynovitis (Spies-Dorgelo et al., 2009)), or included in the guidelines for the treatment of carpal tunnel syndrome (American Association of Orthopaedic Surgeons, 2011), this provides further evidence that a substantial proportion of those with more severe hand problems in this analysis likely had hand OA (81% of those in a problematic baseline hand state had nodes and/ or pain in both hands at baseline). However, the other potential hand conditions described above could also lead to difficulty performing tasks contained within the base model, as well as painful experiences whilst carrying out certain activities. It is also possible that certain painful conditions, such as carpal tunnel syndrome, may be susceptible to quicker recovery times, and potentially reflect individuals more likely to transition between 'least affected' and 'high pain', whilst the high stability of 'severely affected' could reflect individuals most likely to have hand OA. Further research investigating large samples of condition-specific

individuals could explore whether states and transition probabilities identified in this project are generalisable across hand conditions.

11.3.6 Health care data

The analysis of the prescription data highlighted a couple of limitations. Firstly, there was no confirmation that the drug was actually prescribed for a hand problem. A 14 day period following a consultation for hand pain/ problems was permitted to include consulters who received a prescription on the day of consultation, or returned shortly afterwards to collect a prescription. This process followed previous methods performed within the Research Centre (Edwards et al., 2015); however, it is conceivable that the prescription could have been made for a different problem consulted for in that same 14 day period. Secondly, a clear limitation of the analysis of prescription data is that it is unknown whether the prescription was actually dispensed and subsequently taken by the patient. This is a recognised limitation in most studies investigating prescriptions and is difficult to rectify.

As the data collected is from a population-based cohort measured at 3 year intervals, it is difficult to be certain of events (including treatments) that may have occurred to individuals between the assessment time points or to what extent hand symptoms may have fluctuated over the 3 years between measurements. In relation to this, it is difficult to theorise what role treatments may have had on the course of hand problems. Whilst data from medical records was available for the majority of the 5,617 participants analysed, the majority (87%) of these individuals did not access any health care (at least recorded) for hand pain/ problems in the 6 year period. Therefore, it is difficult to assess whether treatments, especially over-the-counter analgesia or other self-management actions may have resulted in improvements in pain symptoms over the 3 year period, with findings demonstrated in *Chapter 6* indicating symptom improvement for some individuals.

11.3.7 Longitudinal assessment of hand state and mental health

This thesis presents the first application of the novel technique of ALTA to MSK conditions (which has been discussed in *section 9.6*). At the time of the project, only three published studies have included the use of ALTA. Furthermore, this thesis is the first study (to my knowledge) to include three fixed time points in the ALTA technique. This analysis has provided insights that would not have been possible with other techniques (such as more restrictive approaches, discussed in *section 11.4.2*). For instance, older people with more severe hand symptoms were more likely to be affected by more severe mental health symptoms, plus a potential relationship of hand pain with depression symptoms, and hand dysfunction with anxiety symptoms was identified. While there were limitations surrounding the application of this technique (discussed in *11.4.1.3*), this exploration has provided hypotheses for future research.

11.4 Latent Transition Analysis

11.4.1 Strengths and limitations

In this study, based on the responses of eight questions, five different phenotypes (profiles) of people with hand pain/ problems were developed through the use of LTA. Thus, a unique feature of LTA is that it contains the strengths and applications of cross-sectional LCA, with a longitudinal approach to classifying individuals into sub-groups based on responses to several indicators of interest.

11.4.1.1 Base model development

LTA can be regarded as a longitudinal extension of the standard LCA which focuses on one time point. A key characteristic of LTA is the derivation of transition probabilities. This represents the estimated probabilities that individuals might change membership of

the identified states over time. While these probabilities can be of interest contextually (for example, which phenotypes are most likely to deteriorate), they also have the flexibility to deal with deterioration versus recovery from a disease. Therefore, LTA can not only reveal longitudinal patterns of symptoms, but also different patterns of disease progression. While there is no universally agreed approach for estimating the required sample size for LTA, generally a sample of 1,000 will generate robust estimates, even with a large number of parameters and a substantial proportion of missing data (Hyatt and Collins, 1998), while about 300 observations is sufficient to perform a basic LTA (Collins and Wugalter, 1992; Collins and Lanza, 2010). As the number of states and time points increases, the required sample size increases. Considering this study used over 5,000 individuals to develop the base model, it can confidently be stated that the results were reliable.

Similarly, there is no 'gold standard' for deciding on the number of states in an LTA model. A general feature of LCA is a 'bootstrap likelihood ratio test', which assesses the current number (n) of latent classes vs. one class less (n-1), to represent whether the current model is more appropriate than a smaller number (Muthén and Muthén, 1998-2015), however this assessment is not (currently) available for LTA. In this project, the standard criteria recommended by Collins and Lanza were used: BIC, entropy and state sample size proportions (Collins and Lanza, 2010). As the AIC criterion has been reported to overestimate the number of classes/ states required in an LCA/ LTA (Yang, 2006; Nylund et al., 2007), this was not used as a criterion for selecting the number of states. Even using the three recommended criteria leaves room for debate with regard to the optimal number of states. In this project, five states were selected as optimum, as this appeared to balance low BIC, high entropy and sufficient sample size proportion in each state. In addition to this, the profiles of each state were investigated to ensure that the interpretation of each state was logical and reflected the hand pain and problems reported

by people assigned to that state, gave sufficiently distinct profiles, and had few indicators with item-response probabilities that do not discriminate well within states (i.e. in the region of 0.4 to 0.6). The six state model was also explored (*section 10.4*), which highlighted that the additional state was a less well-defined sub-state of 'high pain and poor gross function', therefore this was not pursued.

Individuals are placed in states/ phenotypes during the modelling process based on posterior probabilities, with individuals classified in the state with the largest probability, thus there is the potential that individuals may not be clearly allocated to a single state. This may be assessed by the APP for each state, where an average <0.7 would indicate a concern (Clark et al., 2006). The validation of the base model chapter (*Chapter 5*) highlighted the classification of individuals to optimum phenotypes did not appear to be a problem as all APPs were >0.85 *(Table 5.3.1, section 5.3)*. If probabilities were <0.7 for some states, exploring alternative models would be recommended, potentially with a smaller number of states to aid classification.

It was decided in this study to dichotomise the indicators (*Chapter 4*), instead of using the five responses used for the majority of questions. This was enforced to speed up the simulation process, because modelling an LTA method with items of more than two responses increased the processing time exponentially. A potential limitation of this was that the full amount of data was not used, and the states derived could have depended on the cut-off point selected. However, to assess this, a sensitivity analysis was performed (*section 10.2*) where a different cut-point was used for dichotomising indicator responses. Using an alternative dichotomisation at a higher level of severity of pain and function difficulty resulted in a weaker base model due to reduced proportions of individuals considered to have hand pain or function difficulties (*section 10.2*). Testing an alternative cut-point does not resolve the issue of potentially losing information due to

dichotomisation; however it can provide reassurance that other optimal dichotomisations, providing robust findings, have not been missed.

A further limitation of dichotomising the indicator variables is the loss of information regarding the severity of pain and functional difficulties, which is better reflected by the full scale of the items. For example, in the models developed in this study it was not possible to get two states that both represent hand pain, but represent different severities. As a result of dichotomising indictors, the different states can only reflect frequency of indicators (being affected by one, two or three hand pain indicators), rather than variation in severity (for example, being affected by one indicator severely, but another indicator moderately). An approach to incorporate severity would have been to explore a base model development using the original ordinal five response options, or grouping into three (for example, “none or mild”, “moderate”, “severe or extreme”). This would have greatly increased computation time and made interpretation more complex. This is an area for further research.

11.4.1.2 Software

The statistical software of Mplus was used for the development of the base model and further model development. The extensive time needed to perform the simulations was a frequent issue and would often delay progress. In computer simulations exploring a large number of states (particularly six or greater), processing time was usually >24 hours due to the large sample size, three time points with therefore a large number of parameters being estimated. Frequently in these scenarios, the optimum likelihood value was not replicated, and therefore the model had not converged and was unreliable, resulting in the simulation needing to be processed again with a larger number of starting values. In some analyses, 10,000 random start values were reached, resulting in simulation time in excess of 100

hours and the best likelihood was still not replicated; this was reported, where applicable, in the results. Therefore, in these rare cases, there was a potential that the optimal model could have been slightly different. However, during the simulations made throughout the analysis in this thesis, the BIC (and entropy) did not vary by more than 50 (0.01/ 1%) for random start values greater than 500.

Application of LTA is relatively new (especially the extensions used in this thesis such as stratified transition probability analysis, including covariates into the LTA modelling process, and ALTA), therefore technical support was limited as only a few individuals had experience with programming these methods. Support was available from the Mplus developers, however, due to the wide range of analyses Mplus can undertake, expertise on specific queries related to less commonly used procedures (like ALTA) were difficult to come by. Mplus does also offer 'web-notes', which do provide relevant examples for certain queries (in addition to the Mplus User's Guide, Muthén and Muthén, 1998-2015). However, these web-notes can sometimes be technically challenging. One challenge encountered was stratifying transition probabilities by a characteristic such as gender, and computing an additional model to test for a statistical difference (*section 6.4*). Unfortunately, examples of relevant codes to compute the models to perform this assessment were not easily available. Through communication with Dr. Bethany Bray (The Methodology Center, Penn State University), the models were able to be developed successfully, and the code is now available from Dr. Bray to use in potential Mplus/ LTA teaching courses and when other researchers contact The Methodology Center (Penn State) for advice on this specific topic.

Mplus was originally chosen to complete the project analysis mainly because of its flexibility in analysing different forms of LCA, the availability of training courses at the start of the project, and previous informal reports on the flexibility of Mplus. It is possible

to analyse LTA in alternative software. LatentGOLD is a common package used for LCA, and LTA is now possible (labelled 'Latent (Hidden) Markov') within LatentGOLD. In addition to this, WinLTA is a free downloadable package that was originally created by the research team at The Methodology Center (Penn State), however this software is no longer supported by that Centre. A potential reason for this is the team have also developed 'PROC LTA', a procedure accessible in the statistical software SAS. This software, reportedly, has the ability to cope with all the LTA related analysis presented in this project, but currently excluding ALTA.

11.4.1.3 ALTA

A novelty of the ALTA technique used in *Chapter 9* was that longitudinal relationships can be explored in more detail, such as investigating longitudinal hypotheses between two potentially linked domains. However, the disadvantages of working with a novel technique include access to software, and knowledge needed to programme the specific models. ALTA software will hopefully become more accessible for the wider research community in the future.

A limitation of the ALTA work completed in this project was that only the full association model of the ALTA technique was able to be computed. While this model may provide detailed information of the relationship between hand and mental health states, in some scenarios (as mentioned in *section 9.3*) the full association model may not represent the most appropriate relationship. The full association model was, however, theoretically the strongest model to assess the longitudinal associations between hand and mental health states, and conceptually, the full association model was the most appropriate for the research question in *Chapter 9*. Based on the original development of ALTA, without the three other models (independent, cross-sectional and longitudinal models), it was not

possible to declare which was the most appropriate representation of the associations between hand pain/ problems and mental health.

In addition, only individuals with complete data were included. The ALTA software can incorporate individuals with missing data (in a similar way to Mplus), however, the additional algorithm to do this within the software was not available, and the short time line meant that, in this instance at least, individuals with no missing data were the only participants analysed.

A further limitation with the ALTA was the number of parameters estimated, and in conjunction with this, the sample size required. This study used just under 5,000 participants in the analysis, but the effects of estimating 400 values in the η tables were clear, as some probabilities were '0.00' or '1.00', therefore either an empty or full cell. Similar to LTA, there are no 'set number' of participants to include; Flaherty's paper used 6,504 participants, Bray et al. used 1,573, while Witkiewitz and Villarroel only used 872 (Flaherty, 2008; Witkiewitz and Villarroel, 2009; Bray et al., 2010). Therefore, with regard to the ALTA process, a large sample size to determine more reliable estimates is crucial, and with a large number of latent states in each domain, the desired sample size increases exponentially.

11.4.2 Alternative methods

The review of the literature (*Chapter 3*) suggested this was the first time the technique of LTA has been employed to assess MSK conditions with clear clinical hypotheses, as the other two papers were predominately using an adapted LTA technique (*section 3.8*, Reboussin et al., 1999; Von Korff and Miglioretti, 2005). Other studies that have investigated change over time in MSK conditions have used latent class longitudinal methods (LLCA, Dunn et al., 2006; Dunn et al., 2013, and LCGA Dunn et al., 2011;

Nicholls et al., 2014; Rzewuska et al., 2015). LLCA groups individuals into classes reflecting a longitudinal pattern based on response to one repeated indicator measured at multiple time points, whereas LCGA creates longitudinal trajectories based on the response to one repeated factor (generally) where individuals in each trajectory have a similar response profile to other individuals. LCGA incorporates time order of the measurements in the modelling process through the use of linear or quadratic growth functions, while LLCA does not, so LCGA could therefore be deemed more preferable. An extension to LCGA is Growth Mixture Models (GMM) which allows for variation around the mean profiles of the clusters (which is similar to a random effects model). This technique requires a large amount of time to compute, and often provides no additional benefit to LCGA (Strauss et al., 2014). One of the main differences between these latent methods and LTA, is in LTA individuals are permitted to switch or transition between states over the time period analysed, therefore the course and prognosis of a condition (and factors that predict these) is less restrictive in LTA.

Other alternative methods are Generalised Linear Models (GLM), cluster and factor analysis. GLM (Diggle et al., 1994; Agresti et al., 2000) assess the response of one measured factor over a period of follow-up. This method is useful to measure observed data (not latent), and two or more factors could be assessed using an extended multivariate model. However, distinct patterns between the different pain and function indicators, such as those discovered in this project, would be difficult to determine due to the limited number of pain and function factors that realistically can be included. Cluster analysis identifies sub-groups of individuals based on multiple indicator variables. However individuals are automatically allocated to one class only (and assessed at one point in time). Factor analysis identifies clusters of indicators, rather than individuals, and is therefore more suitable for assessing underlying constructs in a variety of indicators.

Alternative approaches to the ALTA technique are dual trajectory LCGA, cross-lagged models, cross-domain growth models and GLMs (mentioned previously, as a bivariate model). Dual trajectory LCGA can also be employed to identify trajectories of two repeatedly measured and related factors, for example pain and function (Olino et al., 2010; Peng, 2011; Twigg, 2014). However, this method is complex, requires a large amount of computation time (occasionally >2 weeks, Peng, 2011), and can still only identify longitudinal patterns based on two items. It could be possible to include additional factors measured over time (more than two) in an LCGA, however, this would naturally increase the complexity and required time. This dual trajectory could potentially be regarded as the equivalent extension to LCGA as ALTA is to LTA.

Cross-lagged models (Burkholder and Harlow, 2003) are longitudinal models in which two variables (domains) predict each other over time. The main difference between this and the ALTA model is that, in ALTA, one variable is regarded as dependent and the research objective is focussed on predicting outcomes in this variable (not a bidirectional relationship). Another crucial difference between the two models is in a cross-lagged model a follow-up outcome of one variable (for example, mental health state) is purely predicted by the other (for example, hand state), but mental health state is not predicted by the change in hand state; an aspect which is contained in the ALTA approach (mental health state at time t is dependent on mental health at time t-1, on hand state at time t-1 and, crucially, at time t).

Cross-domain growth curve models (Sayer and Willett, 1998) assess the relationship between longitudinal trajectories (developed from LCGA). In the ALTA approach, estimates of individual state membership is interpreted (so the probability of membership in states from two domains), however in a cross-domain growth curve model, the correlation between two hypothetically linked trajectories at various time points are

determined. Therefore, the number of estimates from these investigations is relatively small compared to the ALTA approach, and arguably does not provide as much detail about the course of two domains.

Considering the main differences between ALTA and the comparable models (dual trajectory LCGA, cross-lagged model, cross-domain growth curve models), it could be argued that the ALTA technique is more favourable due to the flexibility and freedom of estimation. As mentioned in the original ALTA paper, this additional detail results in a cost of complication (for computing and interpretation), but also reveals "a very detailed picture of developmental change" (Flaherty, 2008). In this study, ALTA was challenging to reproduce and required expert help, but has permitted the exploration of a novel, informative method investigating an additional dimension of hand pain/ problems in older people.

11.5 Further research

Potential research areas, both clinical and methodological, have arisen from the work performed in this project. Firstly, it would be of interest to assess the generalisability of the profiles of hand pain and functional symptoms (external validation); similarly, assess the factors associated with the hand states that have been found in this study, but with a different population to assess the wider applicability of the model. If further exploration of the model revealed similar hand states found in this study, this would suggest that the model is not influenced by sampling variation and is likely to be a suitable representation of hand pain/ problems in other, similar populations. This model, if replicated other populations, could be explored further to investigate whether individuals in particular states may benefit more from specific treatments. For example, the treatment of those in 'high pain' could likely differ to those in 'poor gross function'.

Secondly, radiographic information was not available for all the individuals in this study but there could great value in assessing the relationship of radiographic OA with the different hand states in future research. For example, such analysis could be used to determine if individuals with more severe radiographic features would be classified in 'severely affected', and to investigate if the extent of radiographic OA change predicts future hand state membership. As highlighted in *Chapter 2*, the relationship between symptomatic and radiographic hand OA is uncertain. A potential benefit of this research is that it could highlight whether certain hand states (e.g. hand symptoms) are highly correlated with radiographic subsets (e.g. erosive hand OA, nodal OA). This would potentially develop further understanding of the relationship between symptomatic and radiographic profiles of hand OA.

Thirdly, an improvement in hand symptoms in the older age group was revealed during the course of this thesis which, despite the lower sample size in the older age group, was an important finding. Improvement in hand problems in some older people has been seen previously (Bijsterbosch et al., 2011). One potential explanation for this is acceptance of and adaption to hand pain/ problems and the use of various coping strategies, which has been mentioned in previous work (Myers, 2008; Hill et al., 2011; Bukhave and Huniche, 2014). Further research may further explore these coping strategies and tools that individuals with hand conditions use, in order to increase the knowledge of available options for the wider population. Alternatively, it could be that hand conditions do become less symptomatic in older age, which has previously been suggested (Leung et al., 2013; Prieto-Alhambra et al., 2014). In either case, additional work is required to further understand the relationship between changes in hand symptoms with age in the older population.

Fourthly, an interesting methodological question would be a comparison between the approaches of LTA and other latent variable methodology, such as LCGA. There are differences between the approaches (see *section 11.4.2*), as LCGA classifies individuals according to a single indicator measured repeatedly over time, while LTA classifies individuals based on multiple indicators at a time point and then estimates the probability of transition between states. While the format of these two techniques is generally different, it would be possible to compare the two methodologies for a similar hypothesis, i.e. transition/ trajectories of pain symptoms. A potential methodological study could estimate changes in hand pain over time (assessing the AUSCAN pain score at repeated time points for example) with LCGA and subsequently assess its ability to present a more rounded picture regarding the course and prognosis of hand pain than is achieved via producing hand pain states with LTA (using the AUSCAN pain indicators for example) and permitting participants to move from one state to another. Subsequently, these methods and findings could be presented to clinicians and patients to assess their validity, and investigate which technique is the easiest to understand.

Finally, the technique of ALTA is a complex technique that could be used in many fields and has the ability to assess longitudinal relationships in detail. However, the current availability of the technique is limited due to the complexity of the programming knowledge required to compute it. Therefore, additional work is required to adapt the content of the ALTA software, or develop coding for other statistical packages such as Mplus, to make this more approachable for other researchers.

11.6 Implications of research findings

While replication of findings from this project is desirable in other populations, some key messages can be taken forward for clinicians and other researchers.

Clinicians, in particular GPs who are the primary contact for individuals presenting to health care with hand pain/ problems, should be aware that whilst some individuals with hand pain/ problems will improve, patients with certain characteristics, such as those with poor function, presence of nodes, sleep problems, chronic hand pain duration, bilateral hand pain and symptoms of anxiety/ depression are more likely to have an unfavourable prognosis. These individuals may benefit from additional attention from GPs and may benefit from self-management approaches including occupational therapy, joint protection and advice.

LTA is a valuable method for identifying common relationships between pain and function, and exploring how these change over time. In addition to this, the LTA method has a wide scope of uses within the field of MSK research, where conditions frequently change over time. ALTA can potentially be a powerful and informative longitudinal approach, but further development is needed before reliable and robust results can be produced and presented in health research.

This thesis has led to the development of five phenotypes reflecting various stages of hand pain/ problems in an older population, and explored the patterns of these phenotypes over a 6 year period. Older individuals with hand pain alone were more likely to improve their hand condition, while difficulty with function had a more unfavourable outcome. This work has provided fresh evidence of predictive factors of hand progression, and explored the longitudinal relationship between hand phenotypes and mental health. Finally, this work has applied the statistical technique of LTA, and displayed evidence that this method is a flexible and informative approach for MSK research.

References

Agresti A. Categorical data analysis. New York: Wiley 1990.

Agresti A, Booth JG, Hobert JP, Caffo B. Random effects modeling of categorical response data. Sociol Methodol 2000;30:27-80.

Aihie Sayer A, Poole J, Cox V, et al. Weight from birth to 53 Years; A longitudinal study of the influence on clinical hand osteoarthritis. Arthritis Rheum 2003;48:1030-3.

Akaike H. Factor analysis and AIC. Psychometrika 1987;52:317-32.

Alexander CJ. Heberden's and Bouchard's nodes. Ann Rheum Dis 1999;58:675-8.

Allen KD, Jordan JM, Renner JB, Kraus VB. Relationship of global assessment of change to AUSCAN and pinch and grip strength among individuals with hand osteoarthritis. Osteoarthr Cartil 2006a;14:1281-7.

Allen KD, Jordan JM, Renner JB, Kraus VB. Validity, factor structure, clinical relevance of the AUSCAN osteoarthritis hand index. Arthritis Rheum 2006b;54:551-6.

Altman R, Alarcón G, Appelrouth D, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hand. Arthritis Rheum 1990;33:1601-10.

Altman RD, Dreiser RL, Fisher CL, et al. Diclofenac sodium gel in patients with primary hand osteoarthritis: A randomized, double-blind, placebo-controlled trial. J Rheumatol 2009;36:1991-9.

American Association of Orthopaedic Surgeons 2011. AAOS guideline on the treatment of carpal tunnel syndrome 2011 report for the "re-issue" of the original guideline.

http://www.aaos.org/Research/guidelines/CTS_Treatment_REIssue.pdf. Accessed March 2016.

Arnow BA, Hunkeler EM, Blasey CM, et al. Comorbid depression, chronic pain, and disability in primary care. Psychosom Med 2006;68:262-8.

Arola H-M, Nicholls E, Mallen C, Thomas E. Self-reported pain interference and symptoms of anxiety and depression in community-dwelling older adults: Can a temporal relationship be determined? Eur J Pain 2010;14:966-71.

Ashworth J, Green DJ, Dunn KM, Jordan KP. Opioid use among low back pain patients in primary care: Is opioid prescription associated with disability at 6-month follow-up? Pain 2013;154:1038-44.

Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: A literature review. Arch Intern Med 2003;163:2433-45.

Barrett A, Saava G, Timonen V, Kenny RA. Fifty plus in Ireland 2011. First results from the Irish Longitudinal Study on Ageing (TILDA). www.tilda.ie. Accessed September 2015.

Bedson J, Belcher J, Martino OL, et al. The effectiveness of national guidance in changing analgesic prescribing in primary care from 2002 to 2009: An observational database study. Eur J Pain 2013;17:434-43.

Bellamy N, Kirwan J, Boers M, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. J Rheumatol 1997;24:799-802.

Bellamy N, Campbell J, Haraoui B, et al. Dimensionality and clinical importance of pain and disability in hand osteoarthritis: Development of the Australian/Canadian (AUSCAN) Osteoarthritis Index. Osteoarthr Cartil 2002a;10:855-62.

Bellamy N, Campbell J, Haraoui B, et al. Clinimetric properties of the AUSCAN Osteoarthritis Hand Index: an evaluation of reliability, validity and responsiveness. Osteoarthr Cartil 2002b;10:863-9.

Bellamy N, Wilson C, Hendrikz J. Population-based normative values for the Australian/ Canadian (AUSCAN) hand osteoarthritis index: part 2. Semin Arthritis Rheum 2011;41:149-56.

Bernard TE, Wilder FV, Aluoch M, Leaverton PE. Job-related osteoarthritis of the knee, foot, hand, and cervical spine. J Occup Environ Med 2010;52:33-8.

Bijsterbosch J, Watt I, Meulenbelt I, et al. Clinical and radiographic disease course of hand osteoarthritis and determinants of outcome after 6 years. Ann Rheum Dis 2011;70:68-73.

Bijsterbosch J, Kloppenburg M, Reijnierse M, et al. Association study of candidate genes for the progression of hand osteoarthritis. Osteoarthr Cartil 2013;21:565–9.

Blazer DG. Depression in late life: review and commentary. J Gerontol A Biol Sci Med Sci 2003;58:249-65.

Botha-Scheepers S, Riyazi N, Watt I, et al. Progression of hand osteoarthritis over 2 years: a clinical and radiological follow-up study. Ann Rheum Dis 2009;68:1260-4.

Bowling A. Research Methods in Health: Investigating Health and Health Services. Buckingham: Open University Press 1997.

Bray BC, Lanza ST, Collins LM. Modeling relations among discrete developmental processes: A general approach to associative latent transition analysis. Struct Equ Modeling 2010;17:541-69.

Bukhave EB, Huniche L. Activity problems in everyday life-patients' perspectives of hand osteoarthritis: "try imagining what it would be like having no hands". Disabil Rehabil 2014;36:1636-43.

Bunevicius A, Peceliuniene J, Mickuviene N, Valius L, Bunevicius R. Screening for depression and anxiety disorders in primary care patients. Depress Anxiety. 2007;24:455–60.

Burkholder GJ, Harlow LL. An illustration of a longitudinal cross-lagged design for larger structural equation models. Struct Equ Modeling 2003;10:465-86.

Burton CL, Chesterton LS, Chen Y, van der Windt DA. Clinical course and prognostic factors in conservatively managed carpal tunnel syndrome: A systematic review. Arch Phys Med Rehabil 2015;doi:10.1016/j.apmr.2015.09.013 [Epub ahead of print].

Cain AS, Epler AJ, Steinley D, Sher KJ. Stability and change in patterns of concerns related to eating, weight, and shape in young adult women: a latent transition analysis. J Abnorm Psychol 2010;119:255-67.

Cameron IM, Crawford JR, Lawton K, Reid IC. Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. Br J Gen Pract 2008;58:32-6.

Carmen WJ, Sowers M, Hawthorne VM, Weissfeld LA. Obesity as a risk factor for osteoarthritis of the hand and wrist: a prospective study. Am J Epidemiol 1994;139:119-29.

Caspi D, Flusser G, Farber I, et al. Clinical, radiologic, demographic, and occupational aspects of hand osteoarthritis in the elderly. Semin Arthritis Rheum 2001;30:321-31.

Chaisson CE, Zhang Y, McAlindon TE, et al. Radiographic hand osteoarthritis: incidence, patterns, and influence of pre-existing disease in a population base sample. J Rheumatol 1997;24:1337-43.

Chaisson CE, Zhang Y, Sharma L, Kannel W, Felson DT. Grip strength and the risk of developing radiographic hand osteoarthritis. Results from the Framingham Study. Arthritis Rheum 1999;42:33-8.

Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: a meta-review. World Psychiatry 2014;13:153-60.

Cho S-J, Cohen AS, Kim S-H, Bottge B. Latent transition analysis with a mixture item response theory measurement model. Appl Psych Meas 2010;34:483-504.

Cho S-J, Bottage BA, Cohen AS, Kim S-H. Detecting cognitive change in the math skills of low-achieving adolescents. J Spec Educ 2011;45:67-76.

Chung H, Lanza ST, Loken E. Latent transition analysis: inference and estimation. Stat Med 2008;27:1834-54.

Clark DB, Jones BL, Wood DS, Cornelius JR. Substance use disorder trajectory classes: diachronic integration of onset age, severity, and course. Addict Behav 2006;31:995-1009.

Cleveland MJ, Lanza ST, Ray AE, Turrisi R, Mallett KA. Transitions in first-year college student drinking behaviours: does pre-college drinking moderate the effects of parent- and peer-based intervention components? Psychol Addict Behav 2012;26:440-50.

Collins JE, Katz JN, Dervan EE, Losina E. Trajectories and risk profiles of pain in persons with radiographic, symptomatic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthr Cartil 2014;22:622-30.

Collins LM, Wugalter SE. Latent class models for stage-sequential dynamic latent variables. Multivariate Behav Res 1992;27:131-57.

Collins LM, Lanza ST. Latent class and latent transition analysis: With applications in the social, behavioural, and health sciences. New York: John Wiley & Sons 2010.

Dahaghin S. Hand osteoarthritis: Epidemiology and clinical consequences. Doctoral thesis 2005.

Dahaghin S, Bierma-Zeinstra SMA, Ginai AZ, et al. Prevalence and pattern of radiological hand osteoarthritis and association with pain and disability (The Rotterdam Study). Ann Rheum Dis 2005;64:682-7.

Dahaghin S, Bierma-Zeinstra SMA, Hazes JMW, Koes BW. The clinical burden of radiographic hand osteoarthritis: a systematic appraisal. Arthrit Care Res 2006;55:636-47.

Demissie S, Cupples LA, Myers R, et al. Genome scan for quantity of hand osteoarthritis: the Framingham Study. Arthritis Rheum 2002;46:946-52.

Demyttenaere K, Bruffaerts R, Lee S, et al. Mental disorders among persons with chronic back or neck pain: Results from the world mental health surveys. Pain 2007;129:332-42.

Dieppe P, Cushnaghan J, Tucker M, Browning S, Shepstone L. The Bristol `OA500 study': progression and impact of the disease after 8 years. Osteoarthr Cartil 2000;8:63-8.

Diggle PJ, Liang K-Y, Zeger SL. Analysis of longitudinal data (Vol 13). Oxford, UK: Oxford University Press 1994.

Dillon CF, Hirsch R, Rasch EK, Gu Q. Symptomatic hand osteoarthritis in the United States: prevalence and functional impairment estimates from the Third U.S. National Health and Nutrition Examination Survey. Am J Phys Med Rehabil 2007;86:12-21.

Dominick KL, Jordan JM, Renner JB, Kraus VB. Relationship of radiographic and clinical variables to pinch and grip strength among individuals with osteoarthritis. Arthritis Rheum 2005;52:1424-30.

Dunn KN, Jordan KP, Croft PR. Characterizing the course of low back pain: A Latent Class Analysis. Am J Epidemiol 2006;163:754-61.

Dunn KM, Jordan KP, Mancl L, Drangsholt MR, Le Resche L. Trajectories of pain in adolescents: A prospective cohort study. Pain 2011;152:66-73.

Dunn KM, Campbell P, Jordan KP. Long-term trajectories of back pain: cohort study with 7-year follow-up. BMJ Open 2013;3.

Durrett R. Probability: theory and examples. Cambridge University Press 2004.

Dyer G, Lozano-Calderon S, Gannon C, Baratz M, Ring D. Predictors of acute carpal tunnel syndrome associated with fracture of the distal radius. J Hand Surg Am 2008;33:1309-13.

Dziedzic K, Thomas E, Hill S, et al. The impact of musculoskeletal hand problems in older adults: findings from the North Staffordshire Osteoarthritis Project (NorStOP). Rheumatology (Oxford) 2007;46:963-7.

Dziedzic K. Recent advances in the diagnosis and management of hand osteoarthritis. Int J Clin Rheumatol 2013;8:439-52.

Dziedzic K, Nicholls E, Hill S, et al. Self-management approaches for osteoarthritis in the hand: a 2x2 factorial randomised trial. Ann Rheum Dis 2013;74:108-18.

Edwards JJ, Jordan KP, Peat G, et al. Quality of care for osteoarthritis: the effect of a point-of-care consultation recording template. Rheumatology (Oxford) 2015;54:844-53.

Fautrel B, Hilliquin P, Rozenberg S, et al. Impact of osteoarthritis: results of a nationwide survey of 10,000 patients consulting for OA. Joint Bone Spine 2005;72:235-40.

Felson DT, Couropmitree NN, Chaisson CE, et al. Evidence for a mendelian gene in a segregation analysis of generalized radiographic osteoarthritis. Arthritis Rheum 1998;41:1064-71.

Felson DT, Nevitt, MC. Epidemiologic studies for osteoarthritis: new versus conventional study design approaches. Rheum Dis Clin North Am 2004;30:783-97.

Flaherty BP. Testing the degree of cross-sectional and longitudinal dependence between two discrete dynamic processes. Dev Psychol 2008;44:468-80.

Fiske A, Wetherell JL, Gatz M. Depression in older adults. Annu Rev Clin Psychol 2009;5:363-89.

Geiser C. Data analysis with Mplus. The Guilford Press, New York, New York 2013.

Ghosh B, Gupta M, Mandal S, Ganguly S, Ghosh A. Prevalence and risk factors for hand osteoarthritis. Indian J Rheuma 2014;9:163-6.

Giles JT, Post WS, Blumenthal RS, et al. Longitudinal predictors of progression of carotid atherosclerosis in rheumatoid arthritis. Arthritis Rheum 2011;63:3216-25.

Graham JW, Collins LM, Wugalter SE, Chung NK, Hansen WB. Modeling transitions in latent stage-sequential processes: A substance use prevention example. J Consult Clin Psych 1991;59:48-57.

Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Prevalence and burden of osteoarthritis: results from a population survey in Norway. J Rheumatol 2008a;35:677-84.

Grotle M, Hagen KB, Natvig B, Dahl FA, Kvien TK. Obesity and osteoarthritis in knee, hip and/ or hand: an epidemiological study in the general population with 10 years follow-up. BMC Musculoskelet Disord 2008b;9:132.

Haara MM, Manninen P, Kröger H, et al. Osteoarthritis of finger joints in Finns aged 30 or over: prevalence, determinants, and association with mortality. Ann Rheum Dis 2003;62:151-8.

Hannan MT, Felson DT, Pincus T. Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. J Rheumatol 2000;27:1513-7.

Hart DJ, Spector TD. Cigarette smoking and risk of osteoarthritis in women in the general population: the Chingford study. Ann Rheum Dis 1993;52:93–6.

Haugen IK, Englund M, Aliabadi P, et al. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Ann Rheum Dis 2011a;70:1581-6.

Haugen IK, Moe RH, Slatkowsky-Christensen B, et al. The AUSCAN subscales, AIMS-2 hand/finger subscale, and FIOHA were not unidimensional scales. J Clin Epi 2011b;64:1039-46.

He Y, Zhang M, Lin EH, et al. Mental disorders among persons with arthritis: results from the World Mental Health Surveys. Psychol Med 2008;38:1639-50.

Hill S. Illness perceptions of people with hand problems: A population survey and focus group enquiry. Doctoral thesis, Keele University, UK 2005.

Hill S, Dziedzic K, Thomas E, Baker SR, Croft P. The illness perceptions associated with health and behavioural outcome in people with musculoskeletal hand problems: findings from the North Staffordshire Osteoarthritis Project (NorStOP). Rheumatology (Oxford) 2007;46:944-51.

Hill S, Dziedzic KS, Nio Ong B. Patients' perceptions of the treatment and management of hand osteoarthritis: a focus group enquiry. Disabil Rehabil 2011;33:1866-72.

Hinkle DE, Wiersma W, Jurs SG. Applied statistical for the behavioral sciences. Wadsworth publishing 2003.

Hochberg MC, Altman RD, April KT, et al. American College of Rheumatology 2012 recommendations for the use of nonpharmacologic and pharmacologic therapies in osteoarthritis of the hand, hip, and knee. Arthritis Care Res (Hoboken) 2012;64:465-74.

Holla JFM, van der Leeden M, Heymans MW, et al. Three trajectories of activity limitations in early symptomatic knee osteoarthritis: a 5-year follow-up study. Ann Rheum Dis 2014;73:1369-75.

Hungin AP, Kean WF. Nonsteroidal anti-inflammatory drugs: overused or underused in osteoarthritis? Am J Med 2001;110:8S–11S.

Hunter DJ, Zhang Y, Nevitt MC, et al. Chopstick Arthropathy: the Beijing Osteoarthritis Study. Arthritis Rheum 2004;50:1495-500.

Hyatt SL, Collins LM. Inference in latent transition models: Results from a simulation study. The Methodology Center Technical Report Series #00-38. 1998.

Jenkins CD, Stanton BA, Niemcryk SJ, Rose RM. A scale for the estimation of sleep problems in clinical research. J Clin Epidemiol 1988;41:313-21.

Joint Formulary Committee. Browse BNF. 2015. http://bnf.org/bnf/ (accessed December 01, 2015).

Jones G, Cooley HM, Bellamy N. A cross-sectional study of the association between Heberden's nodes, radiographic osteoarthritis of the hands, grip strength, disability and pain. Osteoarthr Cartil 2001;9:606-11.

Jonsson H, Helgadottir GP, Aspelund T, et al. Hand osteoarthritis in older women is associated with carotid and coronary atherosclerosis: the AGES Reykjavik study. Ann Rheum Dis 2009;68:1696-700.

Jordan K, Clarke AM, Symmons DPM, et al. Measuring disease prevalence: a comparison of musculoskeletal disease using four general practice consultation databases. Br J Gen Pract 2007;57:7-14.

Jordan KP, Kadam UT, Hayward R, et al. Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. BMC Musculoskelet Disord 2010;11:144.

Kalichman L, Kobyliansky E, Malkin I, Yakovenko K, Livshits G. Search for linkage between hand osteoarthritis and 11q 12–13 chromosomal segment. Osteoarthr Cartil 2003;11:561–8.

Kalichman L, Kobyliansky E, Seibel MJ, Livshits G. Repeated measurement study of hand osteoarthritis in an apparently healthy Caucasian population. Am J Hum Biol 2005;17:611-21.

Kalichman L, Kobyliansky E. Age, body composition, and reproductive indices as predictors of radiographic hand osteoarthritis in Chuvashian women. Scan J Rheumatol 2007;36:53-7.

Kalichman L, Hernández-Molina G. Hand osteoarthritis: an epidemiological perspective. Semin Arthritis Rheum 2010;39:465-76.

Kallman AD, Wigley FM, Scott WW Jr, Hochberg MC, Tobin JD. The longitudinal course of hand osteoarthritis in a male population. Arthritis Rheum 1990;33:1323-32.

Kaplan D. An overview of Markov chain methods for the study of stage-sequential developmental processes. Dev Psychol 2008;44:457-67.

Kean WF, Rainsford KD, Kean IR. Management of chronic musculoskeletal pain in the elderly: opinions on oral medication use. Inflammopharmacology 2008;16:53–75.

Kim SK, Park SH, Choe JY. Lower bone mineral density of forearm in postmenopausal patients with radiographic hand osteoarthritis. Rheumatol Int 2010;30:605-12.

Kjeken I, Dagfinrud H, Slatkowsky-Christensen B, et al. Activity limitations and participation restrictions in women with hand osteoarthritis: patients' descriptions and associations between dimensions of functioning. Ann Rheum Dis 2005;64:1633-8.

Kloppenburg M, Stamm T, Watt I, et al. Research in hand osteoarthritis: time for reappraisal and demand for new strategies. An opinion paper. Ann Rheum Dis 2007;66:1157-61.

Kloppenburg M, Kwok WY. Hand osteoarthritis-a heterogeneous disorder. Nat Rev Rheumatol 2011;8:22-31.

Kloppenburg M, Bøyesen P, Smeets W, et al. Report from the OMERACT Hand Osteoarthritis Special Interest Group: advances and future research priorities. J Rheumatol 2014;41:810-8.

Koffeman AR, Van Buul AR, Valkhoff VE, et al. Adverse drug reactions in a primary care population prescribed non-steroidal anti-inflammatory drugs. Scand J Prim Health Care 2015;33:163-9.

Kraus VB, Jordan JM, Doherty M, et al. The Genetics of Generalized Osteoarthritis (GOGO) study: study design and evaluation of osteoarthritis phenotypes. Osteoarthr Cartil 2007;15:120-7.

Kroenke K, Wu J, Bair MJ, et al. Reciprocal relationship between pain and depression: a 12-month longitudinal analysis in primary care. J Pain 2011;12:964-73.

Kwok WY. Clinical aspects of hand osteoarthritis: are erosions of importance? Doctoral thesis, Leiden University, Netherlands 2013.

Lacey RJ, Belcher J, Croft PR. Validity of two simple measures for estimating life-course socio-economic position in cross-sectional postal survey data in an older population: results from the North Staffordshire Osteoarthritis Project (NorStOP). BMC Med Res Methodol 2012;12:88

Lawrence RC, Felson DT, Helmick CG, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part II. Arthritis Rheum 2008;58:26-35.

Lenze EJ, Wetherell JL. A lifespan view of anxiety disorders. Dialogues Clin Neurosci 2011;13:381-99.

Leung GJ, Rainsford KD, Kean WF. Osteoarthritis of the hand I: aetiology and pathogenesis, risk factors, investigation and diagnosis. J Pharm Pharmacol 2013;66:339-46.

Mallen CD, Peat G, Thomas E, Lacey R, Croft P. Predicting poor functional outcome in community-dwelling older adults with knee pain: prognostic value of generic indicators. Ann Rheum Dis 2007;66:1456-61.

Marshall M, Dziedzic KS, van der Windt DA, Hay EM. A systematic search and narrative review of radiographic definitions of hand osteoarthritis in population-based studies. Osteoarthr Cartil 2008;16:219-26.

Marshall M, van der Windt D, Nicholls E, et al. Radiographic hand osteoarthritis: patterns and associations with hand pain and function in a community-dwelling sample. Osteoarthr Cartil 2009;17:1440-7.

Marshall M. Patterns of radiographic hand osteoarthritis and associations with pain and function: a prospective cohort study. Doctoral thesis, Keele University, UK 2010.

Marshall M, Peat G, Nicholls E, et al. Subsets of symptomatic hand osteoarthritis in community-dwelling older adults in the United Kingdom: prevalence, inter-relationships, risk factor profiles and clinical characteristics at baseline and 3-years. Osteoarthr Cartil 2013;21:1674-84.

Massengale M, Reichmann WM, Losina E, Solomon DH, Katz JN. The relationship between hand osteoarthritis and serum leptin concentration in participants of the Third National Health and Nutrition Examination Survey. Arthritis Res Ther 2012;14:R132.

McBeth J, Lacey RJ, Wilkie R. Predictors of new-onset widespread pain in older adults. Results from a population-based prospective cohort study in the UK. Arthritis Rheum 2014;66:757-67.

Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2. The content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. Arthritis Rheum 1992;35:1-10.

Miglioretti DL. Latent transition regression for mixed outcomes. Biometrics 2003;59:710-20.

Muthén LK, Muthén BO. Mplus User's Guide. 7th ed. Los Angeles, CA: Muthén & Muthén 1998-2015.

Muthén B, Muthén LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. Alcohol Clin Exp Res 2000;24:822-91.

Muthén B, Asparouhov T. LTA in Mplus: Transition probabilities influenced by covariates. Mplus Web Notes: No. 13. 2011.

Myers H, Nicholls E, Handy J, et al. The Clinical Assessment Study of the Hand (CAS-HA): a prospective study of musculoskeletal hand problems in the general population. BMC Musculoskelet Disord 2007;8:85.

Myers H. Hand problems and hand function: A clinical epidemiological study in community-dwelling older adults. Doctoral thesis, Keele University, UK 2008.

Nathan PA, Meadows KD, Istvan JA. Predictors of carpal tunnel syndrome: an 11-year study of industrial workers. J Hand Surg Am 2002;27:644-51.

National Institute for Health and Care Excellence (NICE). Osteoarthritis: Care and management in adults. NICE clinical guidelines 177. February 2014.

Nicholls EE, van der Windt DA, Jordan JL, Dziedzic KS, Thomas E. Factors associated with severity and progression of self-reported hand pain and functional difficulty in community-dwelling older adults: a systematic review. Musculoskeletal Care 2012;10:51-62.

Nicholls EE, Thomas E, van der Windt DA, Croft PR, Peat G. Pain trajectory groups in persons with, or at high risk of, knee osteoarthritis: findings from the Knee Clinical Assessment Study and the Osteoarthritis Initiative. Osteoarthr Cartil 2014;22:2041-50.

Núñez E, Steyerberg EW, Núñez J. Regression modeling statigies. Rev Esp Cardiol 2011;64:501-7.

Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. Struct Equ Modeling 2007;14:535-69.

Olino TM, Klein DN, Lewinsohn PM, Rohde P, Seeley JR. Latent trajectory classes of depressive and anxiety disorders from adolescence to adulthood: Descriptions of classes and associations with risk factors. Compr Psychiatry 2010;51:224-35.

Ornstein SM, Nietert PJ, Jenkins RG, Litvin CB. The prevalence of chronic diseases and multimorbidity in primary care practice: A PPRNet report. J Am Board Fam Med 2013;26:518-24.

Østerås N, Hagen KB, Grotle M, et al. Limited effects of exercises in people with hand osteoarthritis: results from a randomized control trial. Osteoarthr Cartil 2014;22:1224-33.

Pallant JF, Bailey CM. Assessment of the structure of the Hospital Anxiety and Depression Scale in musculoskeletal patients. Health Qual Life Outcomes 2005;3:82.

Park M, Unützer J. Geriatric depression in primary care. Psychiatr Clin North Am 2011;34:469-87.

Patrick ME, Collins LM, Smith E, et al. A prospective longitudinal model of substance use onset among South African adolescents. Subst Use Misuse 2009;44:647-62.

Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. Ann Rheum Dis 2001;60:91-7.

Peng VY. Latent trajectories modelling of developmental profiles of multimorbidity over time in primary care. Doctoral thesis, Keele University, UK 2011.

Petras H, Masyn K, Ialongo N. The developmental impact of two first grade preventive interventions on aggressive/disruptive behavior in childhood and adolescence: an application of latent transition growth mixture modeling. Prev Sci 2011;12:300-13.

Prieto-Alhambra D, Judge A, Javaid MK, et al. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: influences of age, gender and osteoarthritis affecting other joints. Ann Rheum Dis 2014;73:1659-64.

Punzi L, Frigato M, Frallonardo P, Ramonda R. Inflammatory osteoarthritis of the hand. Best Pract Res Clin Rheumatol 2010;24:301-12.

Ramaswamy V, Desarbo WS, Reibstein DJ, Robinson WT. An empirical pooling approach for estimating marketing mix elasticities with PIMS data. Market Sci 1993;12:103-24.

Rannou F, Dimet J, Boutron I, et al. Splint for base-of-thumb osteoarthritis: a randomized trial. Ann Intern Med 2009;150:661-9.

Reboussin BA, Liang KY, Reboussin DM. Estimating equations for a latent transition model with multiple discrete indicators. Biometrics 1999;55:839-45.

Richardson JC, Grime JC, Nio Ong B. 'Keeping going': chronic joint pain in older people who describe their health as good. Ageing Soc 2014;34:1380-96.

Riyazi N, Meulenbelt I, Kroon HM, et al. Evidence for familial aggregation of hand, hip, and spine but not knee osteoarthritis in siblings with multiple joint involvement: the GARP study. Ann Rheum Dis 2005;64:438–43.

Roberts ER, Green D, Kadam UT. Chronic condition comorbidity and multidrug therapy in general practice populations: a cross-sectional linkage study. BMJ Open 2014;4.

Rodda J, Walker Z, Carter J. Depression in older adults. BMJ 2011;343.

Rosemann T, Gensichen J, Sauer N, Laux G, Szecsenyi J. The impact of concomitant depression on quality of life and health service utilisation in patients with osteoarthritis. Rheumatol Int 2007;27:859-63.

Rossignol M, Leclerc A, Allaert F, et al. Primary osteoarthritis of hip, knee, and hand in relation to occupational exposure. Occup Environ Med 2005;62:772–7.

Rzewuska M, Mallen CD, Strauss VY, Belcher J, Peat G. One-year trajectories of depression and anxiety symptoms in older patients presenting in general practice with musculoskeletal pain: A latent class growth analysis. J Psychosom Res 2015;79:195-201.

Saunders KW, Dunn KM, Merrill JO, et al. Relationship of opioid use and dosage levels to fractures in older chronic pain patients. J Gen Intern Med 2010;25:310-5.

Sayer AG, Willett JB. A cross-domain model for growth in adolescent alcohol expectancies. Multivar Behav Res 1998;33:509-43.

Schumann A, Meyer C, Rumpf HJ, Hapke U, John U. Naturalistic changes in the readiness to quit tobacco smoking in a German general population sample. Prev Med 2002;35:326-33.

Sclove SL. Application of model-selection criteria to some problems in multivariate analysis. Psychometrika 1987;52:333-43.

Shin SH. A longitudinal examination of the relationships between childhood maltreatment and patterns of adolescent substance use among high-risk adolescents. Am J Addiction 2012;21:453-61.

Spector TD, Cicuttini F, Baker J, Loughlin J, Hard D. Genetic influences on osteoarthritis in women: a twin study. BMJ 1996;312:940-4.

Spies-Dorgelo MN, van der Windt DAWM, van der Horst HE, Prins APA, Stalman WAB. Hand and wrist problems in general practice- patient characteristics and factors related to symptom severity. Rheumatology (Oxford) 2007;46:1723-8.

Spies-Dorgelo MN, van der Windt DA, Prins PA, Dziedzic KS, van der Horst HE. Clinical course and prognosis of hand and wrist problems in primary care. Arthritis Rheum 2008;59:1349-57.

Spies-Dorgelo MN, van der Windt DA, Prins AP, Uitdehaag BM, van der Horst HE. Diagnosis and management of patients with hand and wrist problems in general practice. Eur J Gen Pract 2009;15:84-94.

Stamm TA, Machold KP, Smolen JS, et al. Joint protection and home hand exercises improve hand function in patients with hand osteoarthritis: a randomized controlled trial. Arthritis Rheum 2002;47:44–9.

StataCorp. Stata Statistical Software: Release 13. StataCorp LP; College Station, TX: 2013.

Strauss VY, Jones PW, Kadam UT, Jordan KP. Distinct trajectories of multimorbidity in primary care were identified using latent class growth analysis. J Clin Epidemiol 2014;67:1163-71.

Styrkarsdottir U, Thorleifsson G, Helgadottir HT, et al. Severe osteoarthritis of the hand associates with common variants within the ALDH1A2 gene with rare variants at 1p31. Nat Genet 2014;46:498-502.

Sun G-W, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. J Clin Epidemiol 1996;49:907-16.

Thaper A, Zhang W, Wright G, Doherty M. Relationship between Heberden's nodes and underlying radiographic changes of osteoarthritis. Ann Rheum Dis 2005;64:1214–6.

Thomas E, Wilkie R, Peat G, et al. The North Staffordshire Osteoarthritis Project-NorStOP: Prospective, 3-year study of the epidemiology and management of clinical osteoarthritis in a general population of older adults. BMC Musculoskelet Disord 2004;5:2.

Thomas E, Dunn KM, Mallen C, Peat G. A prognostic approach to defining chronic pain: application to knee pain in older adults. Pain 2008;139:389-97.

Thompson AM, Macy RJ, Fraser MW. Assessing person-centered outcomes in practice research: a latent transition profile framework. J Community Psychol 2011;39:987-1002.

Twigg S. Unravelling the biological, environmental, physical and psychosocial factors impinging on outcomes in early rheumatoid arthritis. Doctoral thesis, University of Leeds, UK 2014.

Uhlig T, Smedstad LM, Vaglum P, et al. The course of rheumatoid arthritis and predictors of psychological, physical and radiographic outcome after 5 years of follow-up. Rheumatology 2000;39:732-41.

van Saase JL, van Romunde LK, Cats A, et al. Epidemiology of osteoarthritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other populations. Ann Rheum Dis 1989;48:271–80.

Verbruggen G, Veys EM. Numerical scoring systems for the anatomic evolution of osteoarthritis of the finger joints. Arthritis Rheum. 1996;39:308-20.

Verkleij SP, Hoekstra T, Rozendaal RM, et al. Defining discriminative pain trajectories in hip osteoarthritis over a 2-year time period. Ann Rheum Dis 2012;71:1517-23.

Vermunt JK, Magidson J. Latent class models for classification. Comput Stat Data An 2003;41:531-7.

Visser AW, den Heijer M, Spoelman W, et al. The association of serum glucose and insulin concentrations with hand osteoarthritis: The NEO study. Osteoarthr Cartil 2013;21:S247.

Visser AW, Ioan-Facsinay A, De Mutsert R, et al. Adiposity and hand osteoarthritis: the Netherlands Epidemiology of Obesity study. Arthritis Res Ther 2014;16:R19.

Von Korff M, Miglioretti DL. A prognostic approach to defining chronic pain. Pain 2005;117:304-13.

Ware J Jr, Kosinski M, Keller SD. A 12-Item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996:34;220-33.

Witkiewitz K, Villarroel NA. Dynamic association between negative affect and alcohol lapses following alcohol treatment. J Consult Clin Psychol 2009;77:633-44.

Wolfe F, Smythe HA, Yunus MB, et al. The American College of Rheumatology 1990 criteria for the classification of Fibromyalgia. Report of the multicentre criteria committee. Arthritis Rheum 1990;33:160-72.

Wolitzky-Taylor KB, Castriotta N, Lenze EJ, Stanley MA, Craske MG. Anxiety disorders in older adults: a comprehensive review. Depress Anxiety 2010;27:190-211.

Yang C-C. Evaluating latent class analysis models in qualitative phenotype identification. Comput Stat Data An 2006;50:1090-104.

Yeşil H, Hepgüler S, Öztürk C, Yeşil M, Çapaci K. Risk factors of symptomatic knee, hand and hip osteoarthritis in a suburban area of İzmir city. Turk J Phys Med Rehab 2014;60:126-33.

Yusef E, Nelissen RG, Ioan-Facsinay A, et al. Association between weight or body mass index and hand osteoarthritis: a systematic review. Ann Rheum Dis 2010;69:761-5.

Zhang W, Doherty M, Leeb BF, et al. EULAR evidence based recommendations for the management of hand osteoarthritis: report of a Task Force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). Ann Rheum Dis 2007;66:377-88.

Zhang W, Doherty M, Leeb BF, et al. EULAR evidence-based recommendations for the diagnosis of hand osteoarthritis: report of a task force of the ESCISIT. Ann Rheum Dis 2009;68:8-17.

Zhang Y, Niu J, Kelly-Hayes M, et al. Prevalence of symptomatic hand osteoarthritis and its impact on functional status among the elderly: The Framingham Study. Am J Epidemiol 2002;156:1021-7.

Zhang Y, Xu L, Nevitt MC, et al. Lower prevalence of hand osteoarthritis among Chinese subjects in Beijing compared with white subjects in the United States. Arthritis Rheum 2003;48:1034-40.

Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361-70.

Appendices

Appendix A: The AUSCAN questionnaire

The following questions concern the amount of pain you have experienced in your hands.

For each situation please enter the amount of **pain** experienced in the **last week**.

(Please put a cross in one box on each line)

QUESTION: How much pain do you have in your hands?

a. At rest (i.e. when not using your hands).

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

b. When gripping objects with your hands.

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

c. When lifting objects with your hands.

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

d. When turning objects with your hands.

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

e. When squeezing objects with your hands.

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

The following question concerns the amount of **joint stiffness** (not pain) you have experienced **in the last week** in your **hands**. Stiffness is a sensation of restriction or slowness in the ease with which you move your hands.

(Please put a cross in the box of the most appropriate answer)

How **severe** is stiffness in your hands **after first wakening** in the morning?

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

The following questions concern your **physical function**. By this we mean your ability to move around and to look after yourself. For each of the following activities, please indicate the **degree of difficulty** you have experienced **in the last week** due to your **hand problem.**

(Please put a cross in one box on each line)

QUESTION: How much difficulty do you have with the following?

a. Turning taps on.

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

b. Turning a round door-knob or handle.

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

c. Doing up buttons.

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

QUESTION: How much difficulty do you have with the following?

(Please put a cross in one box on each line)

d. Fastening jewellery (e.g. watches, earrings, cufflinks, necklaces, brooches and bracelets).

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

e. Opening a new jar.

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

f. Carrying a full pot with one hand (e.g. carrying any reasonably heavy object such as a saucepan).

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

g. Peeling vegetables/fruits.

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

h. Picking up large heavy objects.

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

i. Wringing out washcloths (e.g. squeezing a wet sponge or flannel).

None	Mild	Moderate	Severe	Extreme
☐	☐	☐	☐	☐

Appendix B: RUG meeting indicators

AUSCAN (15 items):

- How much pain do you have:
  - At rest?
- How much pain do you have:
  - Gripping objects?
- How much pain do you have:
  - When lifting objects?
- How much pain do you have:
  - When turning objects with your hands?
- How much pain do you have:
  - When squeezing objects with hands?
- How severe is stiffness in hands after first waking in the morning?
- How much difficulty do you have:
  - Turning taps on?
- How much difficulty do you have:
  - Turning a door-knob or handle?
- How much difficulty do you have:
  - Doing up buttons?
- How much difficulty do you have:
  - Fastening jewellery?
- How much difficulty do you have:
  - Opening a new jar?
- How much difficulty do you have:

    - Carrying a full pot with one hand (i.e. a saucepan)?
- How much difficulty do you have:
    - Peeling vegetables/ fruit?
- How much difficulty do you have:
    - Picking up large heavy objects?
- How much difficulty do you have:
    - Wringing out washcloths (i.e. a flannel)?

AIMS2 (Hand and Finger Function) (3 items):

- Could you easily write with a pen or pencil?
- Could you easily turn a key in a lock?
- Could you easily tie a knot/bow?

AIMS2 (Arthritis Pain) (3 items)

- How often did you have pain in two or more hand joints?
- How often did morning stiffness in your hands last more than one hour from waking up?
- How did your hand problems make it difficult for you to sleep?

AIMS2 (Overall Arthritis Impact) (1 item)

- How often have you taken medication for hand symptoms?

Stand-alone questions (12 items):

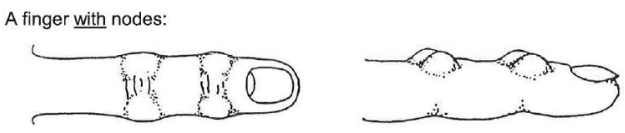
- Hand problems in both hands (or one hand only)
- How much hand stiffness do you usually have?
- How much hand aching do you usually have?
- How much hand tenderness do you usually have?
- How much hand weakness do you usually have?
- How much hand clumsiness do you usually have?
- How much burning sensation do you usually have in your hand?
- How much hand tingling do you usually have?
- How much hand numbness do you usually have?
- How often did your hands feel hot or warm?
- How often did your hand problems make you feel frustrated?
- How often did hand problems cause you to drop objects?

Previous hand experiences (5 items):

- Duration of hand pain
- Previous injury to hand
- Operation on hand
- Job that has involved excessive use of hands
- Hobby/ pastime that has involved excessive use of hands

Presence of nodes (1 item):

- Finger(s) with nodes

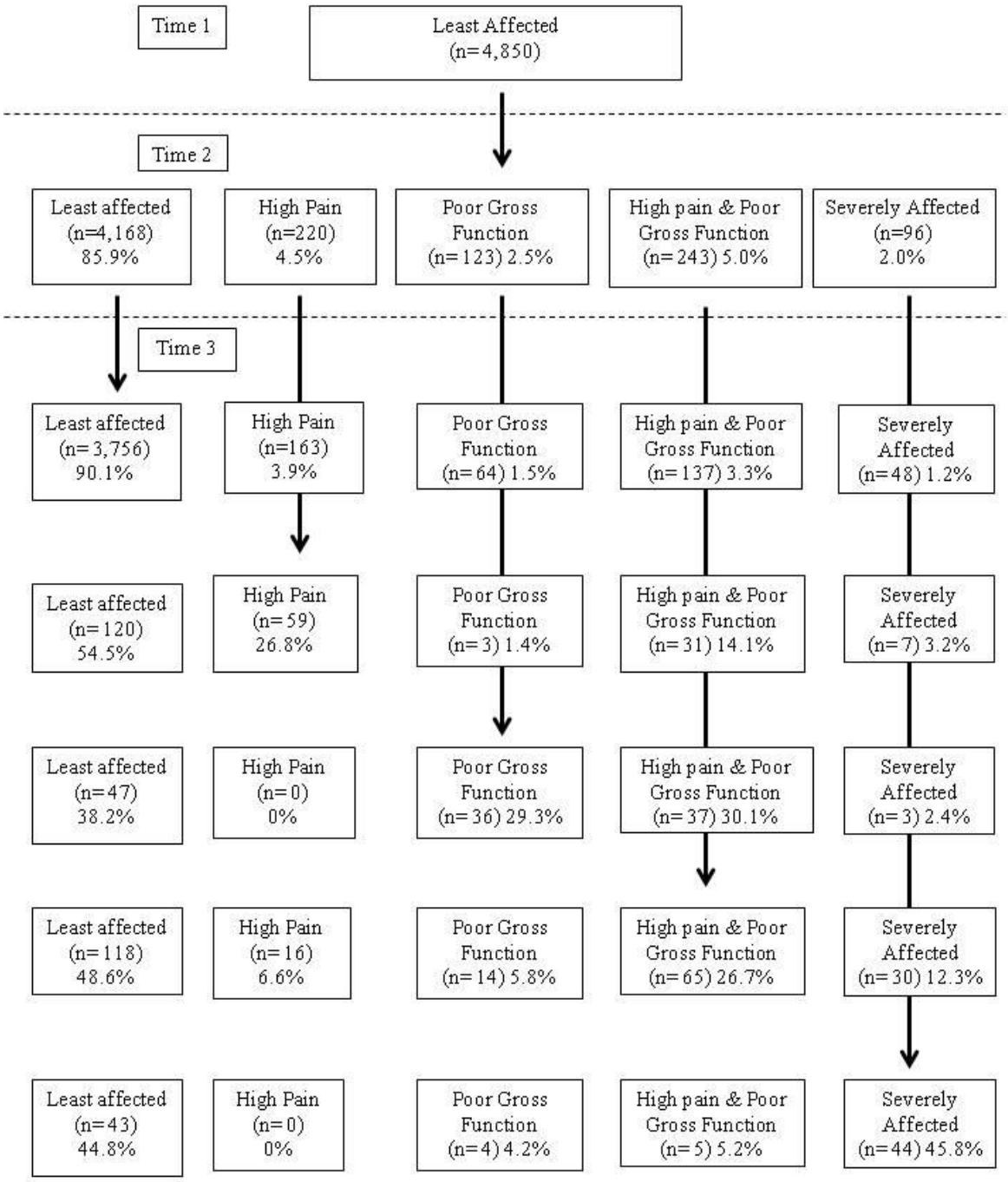


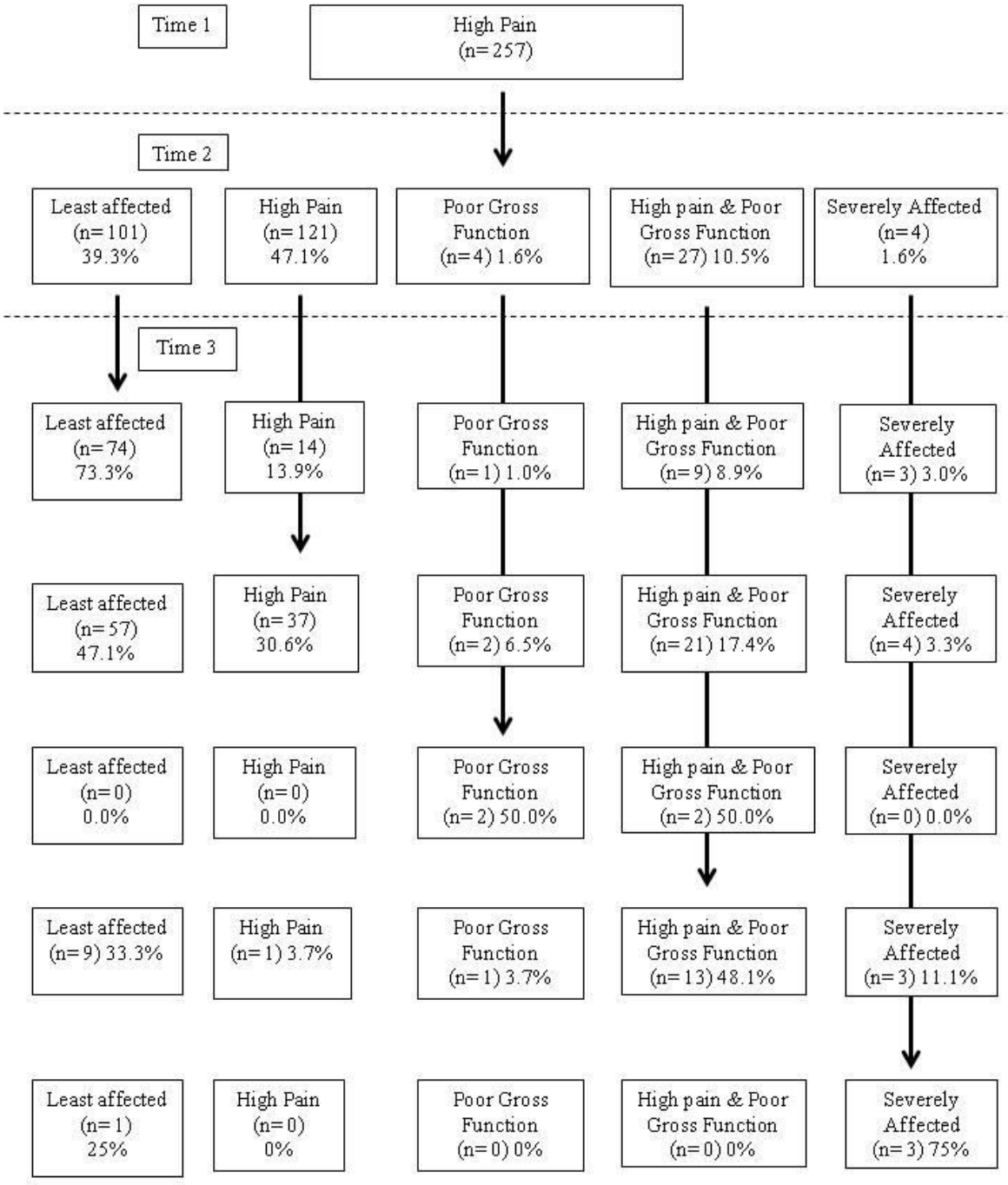
Appendix C: 3 year pilot LTA

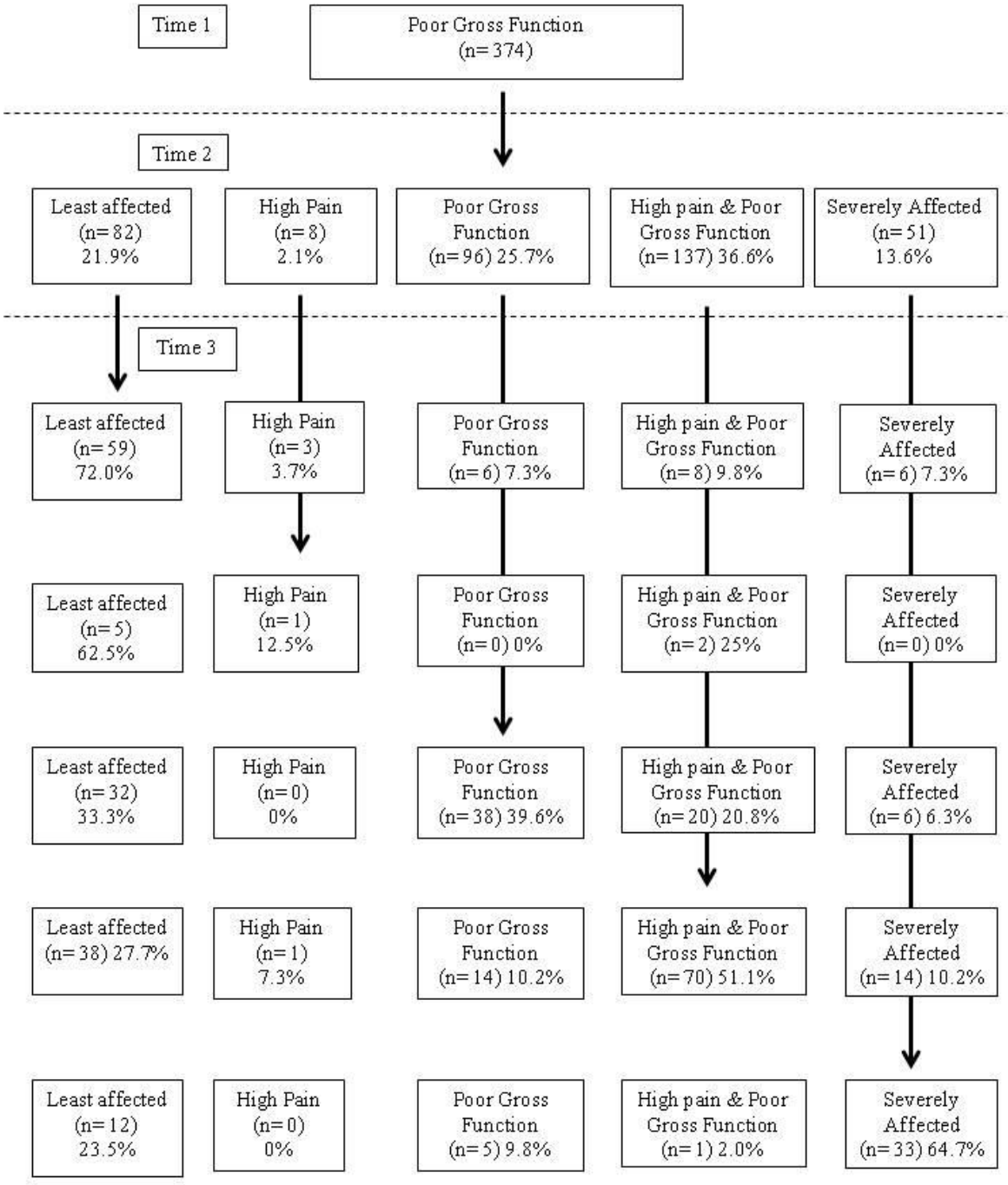
Appendix C.1: LTA model using participant who responded at baseline and 3 years.

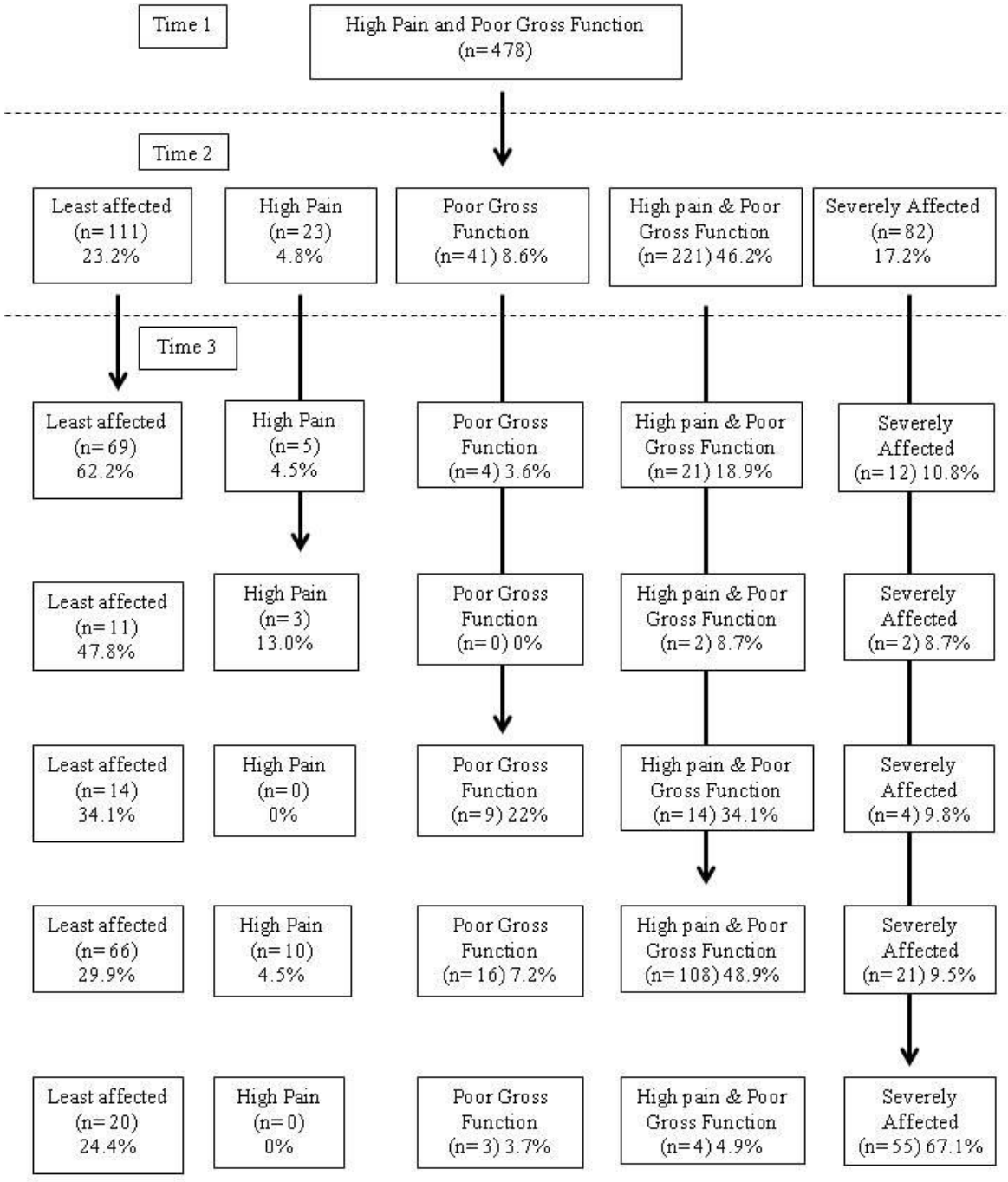
	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Latent state proportions					
Baseline (Time 1)	0.746	0.045	0.056	0.069	0.084
3 years (Time 2)	0.706	0.056	0.042	0.095	0.101
Item-response probabilities					
Pain when turning objects	0.000	0.671	0.130	0.889	0.973
Pain when squeezing objects	0.003	0.817	0.207	0.950	0.990
Pain when gripping objects	0.005	0.708	0.140	0.844	0.954
Difficulty opening a new jar	0.008	0.220	0.754	0.883	1.000
Difficulty carrying a full pot	0.006	0.116	0.695	0.840	0.988
Difficulty wringing out a dishcloth	0.003	0.200	0.551	0.785	0.990
Difficulty doing-up buttons	0.001	0.044	0.169	0.212	0.909
Difficulty turning taps on	0.000	0.012	0.152	0.148	0.914
Latent transition probabilities					
Baseline to 3 years	Least affected	High pain	Poor gross function	High pain & poor gross function	Severely affected
Least affected	0.866	0.043	0.025	0.042	0.025
High pain	0.471	0.371	0.029	0.111	0.018
Poor gross function	0.265	0.024	0.224	0.344	0.143
High pain & poor gross function	0.187	0.067	0.100	0.427	0.219
Severely affected	0.135	0.038	0.010	0.120	0.697

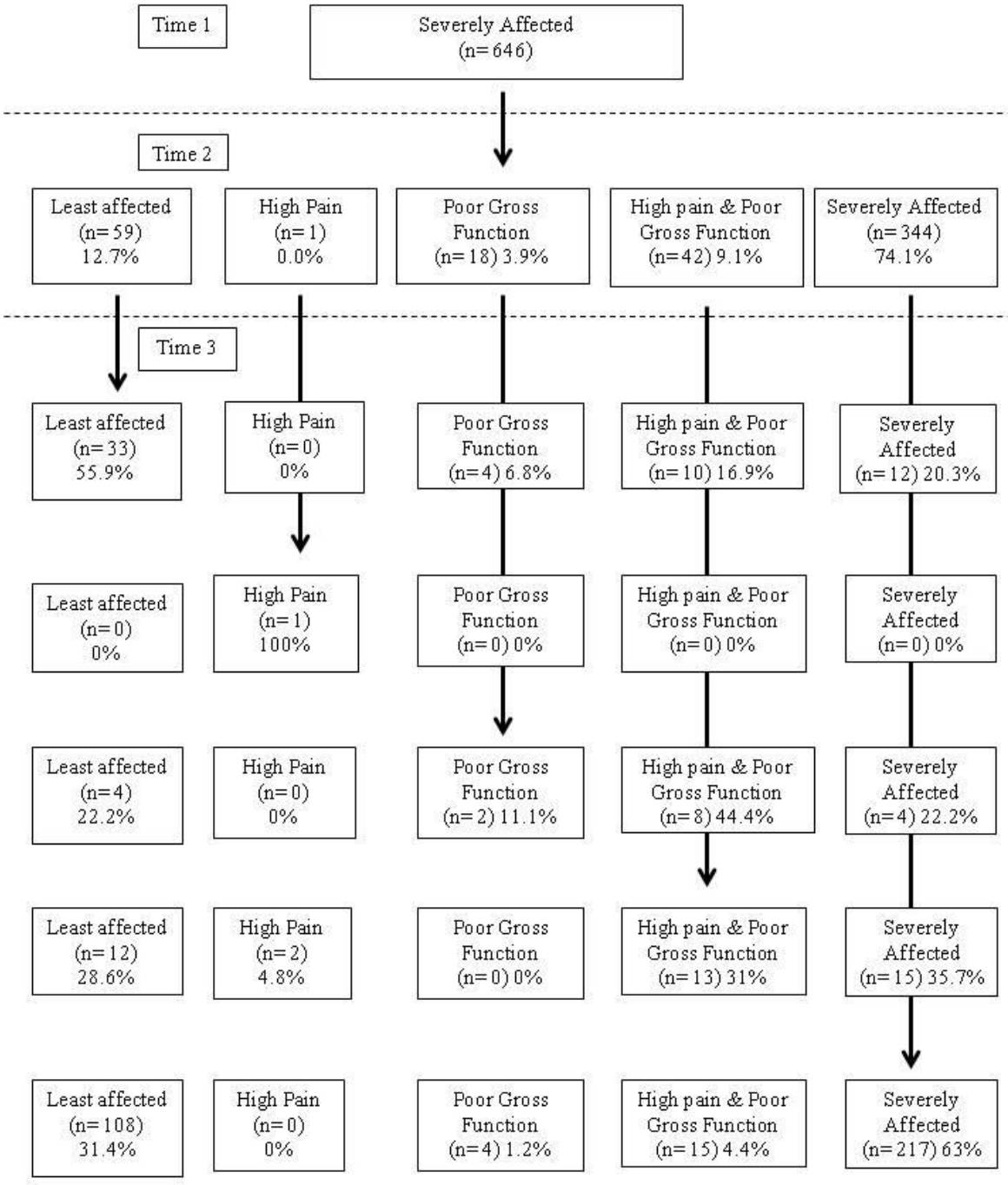
Appendix D: 6 Year transition pathways




Time 1
High Pain (n=257)
Time 2
Least affected (n=101) 39.3%
High Pain (n=121) 47.1%
Poor Gross Function (n=4) 1.6%
High pain & Poor Gross Function (n=27) 10.5%
Severely Affected (n=4) 1.6%
Time 3
Least affected (n=74) 73.3%
High Pain (n=14) 13.9%
Poor Gross Function (n=1) 1.0%
High pain & Poor Gross Function (n=9) 8.9%
Severely Affected (n=3) 3.0%
Least affected (n=57) 47.1%
High Pain (n=37) 30.6%
Poor Gross Function (n=2) 6.5%
High pain & Poor Gross Function (n=21) 17.4%
Severely Affected (n=4) 3.3%
Least affected (n=0) 0.0%
High Pain (n=0) 0.0%
Poor Gross Function (n=2) 50.0%
High pain & Poor Gross Function (n=2) 50.0%
Severely Affected (n=0) 0.0%
Least affected (n=9) 33.3%
High Pain (n=1) 3.7%
Poor Gross Function (n=1) 3.7%
High pain & Poor Gross Function (n=13) 48.1%
Severely Affected (n=3) 11.1%
Least affected (n=1) 25%
High Pain (n=0) 0%
Poor Gross Function (n=0) 0%
High pain & Poor Gross Function (n=0) 0%
Severely Affected (n=3) 75%


Time 1
Poor Gross Function (n=374)
Time 2
Least affected (n=82) 21.9%
High Pain (n=8) 2.1%
Poor Gross Function (n=96) 25.7%
High pain & Poor Gross Function (n=137) 36.6%
Severely Affected (n=51) 13.6%
Time 3
Least affected (n=59) 72.0%
High Pain (n=3) 3.7%
Poor Gross Function (n=6) 7.3%
High pain & Poor Gross Function (n=8) 9.8%
Severely Affected (n=6) 7.3%
Least affected (n=5) 62.5%
High Pain (n=1) 12.5%
Poor Gross Function (n=0) 0%
High pain & Poor Gross Function (n=2) 25%
Severely Affected (n=0) 0%
Least affected (n=32) 33.3%
High Pain (n=0) 0%
Poor Gross Function (n=38) 39.6%
High pain & Poor Gross Function (n=20) 20.8%
Severely Affected (n=6) 6.3%
Least affected (n=38) 27.7%
High Pain (n=1) 7.3%
Poor Gross Function (n=14) 10.2%
High pain & Poor Gross Function (n=70) 51.1%
Severely Affected (n=14) 10.2%
Least affected (n=12) 23.5%
High Pain (n=0) 0%
Poor Gross Function (n=5) 9.8%
High pain & Poor Gross Function (n=1) 2.0%
Severely Affected (n=33) 64.7%


Time 1
High Pain and Poor Gross Function (n=478)
Time 2
Least affected (n=111) 23.2%
High Pain (n=23) 4.8%
Poor Gross Function (n=41) 8.6%
High pain & Poor Gross Function (n=221) 46.2%
Severely Affected (n=82) 17.2%
Time 3
Least affected (n=69) 62.2%
High Pain (n=5) 4.5%
Poor Gross Function (n=4) 3.6%
High pain & Poor Gross Function (n=21) 18.9%
Severely Affected (n=12) 10.8%
Least affected (n=11) 47.8%
High Pain (n=3) 13.0%
Poor Gross Function (n=0) 0%
High pain & Poor Gross Function (n=2) 8.7%
Severely Affected (n=2) 8.7%
Least affected (n=14) 34.1%
High Pain (n=0) 0%
Poor Gross Function (n=9) 22%
High pain & Poor Gross Function (n=14) 34.1%
Severely Affected (n=4) 9.8%
Least affected (n=66) 29.9%
High Pain (n=10) 4.5%
Poor Gross Function (n=16) 7.2%
High pain & Poor Gross Function (n=108) 48.9%
Severely Affected (n=21) 9.5%
Least affected (n=20) 24.4%
High Pain (n=0) 0%
Poor Gross Function (n=3) 3.7%
High pain & Poor Gross Function (n=4) 4.9%
Severely Affected (n=55) 67.1%


Time 1
Severely Affected (n=646)
Time 2
Least affected (n=59) 12.7%
High Pain (n=1) 0.0%
Poor Gross Function (n=18) 3.9%
High pain & Poor Gross Function (n=42) 9.1%
Severely Affected (n=344) 74.1%
Time 3
Least affected (n=33) 55.9%
High Pain (n=0) 0%
Poor Gross Function (n=4) 6.8%
High pain & Poor Gross Function (n=10) 16.9%
Severely Affected (n=12) 20.3%
Least affected (n=0) 0%
High Pain (n=1) 100%
Poor Gross Function (n=0) 0%
High pain & Poor Gross Function (n=0) 0%
Severely Affected (n=0) 0%
Least affected (n=4) 22.2%
High Pain (n=0) 0%
Poor Gross Function (n=2) 11.1%
High pain & Poor Gross Function (n=8) 44.4%
Severely Affected (n=4) 22.2%
Least affected (n=12) 28.6%
High Pain (n=2) 4.8%
Poor Gross Function (n=0) 0%
High pain & Poor Gross Function (n=13) 31%
Severely Affected (n=15) 35.7%
Least affected (n=108) 31.4%
High Pain (n=0) 0%
Poor Gross Function (n=4) 1.2%
High pain & Poor Gross Function (n=15) 4.4%
Severely Affected (n=217) 63%

Appendix E: LTA with covariates

Appendix E.1: Latent Transition Analysis parameters when using gender as a covariate

	LA	HP	PGF	HPPGF	SA
Latent state proportions					
Baseline (Time 1)	0.767	0.045	0.059	0.067	0.062
3 years (Time 2)	0.721	0.060	0.048	0.094	0.079
6 years (Time 3)	0.702	0.058	0.047	0.093	0.100
Item-response probabilities					
Pain when turning objects	0.001	0.739	0.125	0.913	0.977
Pain when squeezing objects	0.004	0.825	0.152	0.959	0.988
Pain when gripping objects	0.006	0.777	0.139	0.862	0.973
Difficulty opening a new jar	0.004	0.234	0.724	0.905	1.000
Difficulty carrying a full pot	0.004	0.093	0.623	0.833	0.993
Difficulty wringing out a dishcloth	0.002	0.195	0.434	0.790	0.988
Difficulty doing-up buttons	0.001	0.047	0.159	0.242	0.918
Difficulty turning taps on	0.000	0.016	0.090	0.165	0.892
Latent transition probabilities					
Baseline to 3 years	LA	HP	PGF	HPPGF	SA
Least affected	0.867	0.048	0.027	0.040	0.018
High pain	0.406	0.383	0.035	0.148	0.028
Poor gross function	0.262	0.036	0.270	0.318	0.114
High pain & PGF	0.204	0.044	0.090	0.451	0.211
Severely affected	0.135	0.009	0.057	0.116	0.684
3 years to 6 years	LA	HP	PGF	HPPGF	SA
Least affected	0.867	0.048	0.024	0.038	0.023
High pain	0.465	0.260	0.037	0.176	0.063
Poor gross function	0.300	0.005	0.347	0.266	0.083
High pain & PGF	0.221	0.080	0.089	0.413	0.196
Severely affected	0.176	0.000	0.034	0.057	0.734

Appendix E.2: Latent Transition Analysis parameters when using age as a covariate.

	LA	HP	PGF	HPPGF	SA
Latent state proportions					
Baseline (Time 1)	0.768	0.044	0.058	0.067	0.063
3 Years (Time 2)	0.721	0.059	0.046	0.094	0.080
6 Years (Time 3)	0.702	0.058	0.045	0.093	0.101
Item-response probabilities					
Pain when turning objects	0.001	0.731	0.124	0.916	0.976
Pain when squeezing objects	0.004	0.813	0.158	0.960	0.989
Pain when gripping objects	0.006	0.758	0.147	0.866	0.973
Difficulty opening a new jar	0.005	0.228	0.735	0.897	1.000
Difficulty carrying a full pot	0.005	0.092	0.637	0.819	0.993
Difficulty wringing out a dishcloth	0.002	0.180	0.448	0.786	0.987
Difficulty doing-up buttons	0.001	0.038	0.175	0.232	0.917
Difficulty turning taps on	0.000	0.014	0.095	0.159	0.886
Latent transition probabilities					
Baseline to 3 years	LA	HP	PGF	HPPGF	SA
Least affected	0.867	0.049	0.025	0.040	0.019
High pain	0.416	0.385	0.027	0.151	0.022
Poor gross function	0.245	0.029	0.272	0.324	0.130
High pain & PGF	0.206	0.037	0.093	0.454	0.210
Severely affected	0.136	0.006	0.059	0.118	0.682
3 years to 6 years	LA	HP	PGF	HPPGF	SA
Least affected	0.868	0.049	0.023	0.037	0.023
High pain	0.482	0.262	0.029	0.173	0.054
Poor gross function	0.286	0.000	0.350	0.270	0.094
High pain & PGF	0.219	0.077	0.089	0.418	0.197
Severely affected	0.178	0.001	0.033	0.056	0.733

Appendix E.3: Latent Transition Analysis parameters when using presence widespread pain as a covariate.

	LA	HP	PGF	HPPGF	SA
Latent state proportions					
Baseline (Time 1)	0.767	0.044	0.059	0.068	0.063
3 years (Time 2)	0.721	0.059	0.047	0.094	0.079
6 years (Time 3)	0.705	0.059	0.046	0.093	0.097
Item-response probabilities					
Pain when turning objects	0.001	0.735	0.124	0.915	0.976
Pain when squeezing objects	0.004	0.819	0.154	0.960	0.989
Pain when gripping objects	0.006	0.765	0.146	0.864	0.973
Difficulty opening a new jar	0.005	0.228	0.723	0.898	1.000
Difficulty carrying a full pot	0.004	0.090	0.626	0.822	0.993
Difficulty wringing out a dishcloth	0.002	0.182	0.441	0.787	0.988
Difficulty doing-up buttons	0.001	0.039	0.169	0.238	0.917
Difficulty turning taps on	0.000	0.013	0.091	0.162	0.889
Latent transition probabilities					
Baseline to 3 years	LA	HP	PGF	HPPGF	SA
Least affected	0.867	0.048	0.026	0.040	0.019
High pain	0.422	0.380	0.026	0.151	0.020
Poor gross function	0.251	0.033	0.273	0.324	0.119
High pain & PGF	0.207	0.037	0.093	0.450	0.213
Severely affected	0.134	0.007	0.058	0.118	0.683
3 years to 6 years	LA	HP	PGF	HPPGF	SA
Least affected	0.861	0.050	0.024	0.041	0.025
High pain	0.500	0.258	0.031	0.159	0.051
Poor gross function	0.307	0.003	0.349	0.259	0.083
High pain & PGF	0.246	0.080	0.089	0.399	0.186
Severely affected	0.224	0.000	0.035	0.055	0.686

Appendix E.4: Latent Transition Analysis parameters when using living status as a covariate.

	LA	HP	PGF	HPPGF	SA
Latent state proportions					
Baseline (Time 1)	0.769	0.043	0.058	0.069	0.061
3 years (Time 2)	0.721	0.061	0.047	0.095	0.076
6 years (Time 3)	0.703	0.058	0.046	0.095	0.098
Item-response probabilities					
Pain when turning objects	0.001	0.731	0.127	0.912	0.979
Pain when squeezing objects	0.004	0.822	0.152	0.959	0.990
Pain when gripping objects	0.006	0.762	0.145	0.867	0.974
Difficulty opening a new jar	0.005	0.229	0.726	0.901	1.000
Difficulty carrying a full pot	0.005	0.095	0.635	0.821	0.992
Difficulty wringing out a dishcloth	0.002	0.183	0.447	0.786	0.987
Difficulty doing-up buttons	0.001	0.035	0.168	0.243	0.915
Difficulty turning taps on	0.000	0.014	0.093	0.163	0.891
Latent transition probabilities					
Baseline to 3 years	LA	HP	PGF	HPPGF	SA
Least affected	0.868	0.050	0.026	0.039	0.016
High pain	0.403	0.387	0.024	0.161	0.025
Poor gross function	0.241	0.033	0.278	0.328	0.120
High pain & PGF	0.207	0.040	0.092	0.455	0.207
Severely affected	0.133	0.006	0.055	0.125	0.681
3 years to 6 years	LA	HP	PGF	HPPGF	SA
Least affected	0.867	0.049	0.023	0.038	0.023
High pain	0.482	0.257	0.032	0.179	0.050
Poor gross function	0.294	0.000	0.355	0.270	0.082
High pain & PGF	0.220	0.079	0.089	0.415	0.197
Severely affected	0.174	0.000	0.033	0.056	0.737

Appendix F: Frequency of predictive factors

Appendix F.1: Frequency (proportion (%'s)) of individuals in each hand state at baseline (bold), 3 years (underlined) and 6 years (standard font) for predictors in group one.

Baseline factor		LA	HP	PGF	HPPGF	SA
Gender	Male	N.A.	**128 (57.1%)**	**66 (21.5%)**	**128 (32.5%)**	**93 (26.3%)**
		112 (38.5%)	69 (58.5%)	28 (19.9%)	112 (31.6%)	94 (25.1%)
		145 (39.6%)	49 (62.0%)	33 (29.2%)	96 (29.5%)	92 (23.3%)
Age (years)	50-64	N.A.	**147 (65.6%)**	**160 (52.1%)**	**235 (59.6%)**	**176 (49.7%)**
		146 (50.2%)	84 (71.2%)	80 (56.7%)	219 (61.9%)	189 (50.4%)
		193 (52.7%)	60 (76.0%)	62 (54.9%)	201 (61.7%)	202 (51.1%)
	65-74	N.A.	**60 (26.8%)**	**110 (35.8%)**	**124 (31.5%)**	**134 (37.9%)**
		105 (36.1%)	32 (27.1%)	45 (31.9%)	110 (31.1%)	136 (36.3%)
		123 (33.6%)	16 (20.3%)	42 (37.2%)	103 (31.6%)	144 (34.5%)
	75+	N.A.	**17 (7.6%)**	**37 (12.1%)**	**35 (8.9%)**	**44 (12.4%)**
		40 (13.8%)	2 (1.7%)	16 (11.4%)	25 (7.1%)	50 (13.3%)
		50 (13.7%)	3 (3.8%)	9 (8.0%)	22 (6.8%)	49 (12.4%)
Lived Alone	Yes	N.A.	**29 (13.6%)**	**60 (20.1%)**	**85 (22.1%)**	**94 (28.3%)**
		61 (22.1%)	14 (12.2%)	29 (21.3%)	70 (20.2%)	94 (26.4%)
		78 (22.2%)	11 (14.1%)	19 (17.1%)	68 (21.4%)	92 (24.8%)

Baseline factor		LA	HP	PGF	HPPGF	SA
Employment Status	Retired	N.A.	**96 (43.4%)**	**163 (54.5%)**	**188 (49.3%)**	**187 (55.3%)**
		159 (56.6%)	38 (33.0%)	75 (54.7%)	158 (45.7%)	204 (56.7%)
		194 (54.3%)	26 (32.9%)	60 (54.6%)	144 (45.4%)	210 (55.9%)
	Employed	N.A.	**92 (41.6%)**	**78 (26.1%)**	**88 (23.1%)**	**44 (13.0%)**
		63 (22.4%)	56 (48.7%)	36 (26.3%)	99 (28.6%)	48 (13.3%)
		85 (23.8%)	37 (46.8%)	30 (27.3%)	92 (29.0%)	58 (15.4%)
	Other	N.A.	**33 (14.9%)**	**58 (19.4%)**	**105 (27.6%)**	**107 (31.7%)**
		59 (21.0%)	21 (18.3%)	26 (19.0%)	89 (25.7%)	108 (30.0%)
		78 (21.9%)	16 (20.3%)	20 (18.2%)	81 (25.6%)	108 (28.7%)
Social Class	Higher managerial/ Professional	N.A.	**66 (30.3%)**	**75 (25.7%)**	**77 (20.6%)**	**51 (15.9%)**
		55 (20.5%)	37 (32.5%)	33 (24.4%)	80 (23.5%)	64 (18.6%)
		67 (19.6%)	23 (29.5%)	33 (29.5%)	78 (25.1%)	68 (18.9%)
	Intermediate	N.A.	**52 (23.9%)**	**81 (27.7%)**	**94 (25.2%)**	**78 (24.4%)**
		75 (27.9%)	30 (26.3%)	29 (21.5%)	83 (24.3%)	88 (25.6%)
		93 (27.2%)	21 (26.9%)	26 (23.1%)	85 (27.3%)	80 (22.2%)
	Routine/ Manual	N.A.	**100 (45.9%)**	**136 (46.6%)**	**202 (54.2%)**	**191 (56.7%)**
		139 (51.7%)	47 (41.2%)	73 (54.1%)	178 (52.2%)	192 (55.8%)
		182 (53.2%)	34 (43.6%)	53 (47.3%)	148 (47.6%)	212 (58.9%)

Appendix F.2: Frequency (proportion (%'s) unless stated) of individuals in each hand state at baseline (bold), 3 years (underlined) and 6 years (standard font) for predictors in group two.

Baseline factor		LA	HP	PGF	HPPGF	SA
ACR Widespread pain		N.A.	**105 (46.9%)**	**141 (45.9%)**	**227 (57.6%)**	**246 (69.5%)**
		124 (42.6%)	60 (50.9%)	68 (48.2%)	213 (60.2%)	254 (67.7%)
		179 (48.9%)	44 (55.7%)	59 (52.2%)	180 (55.2%)	257 (65.1%)
HADS	Depression (n (mean))	N.A.	**222 (3.88)**	**301 (4.63)**	**390 (5.25)**	**346 (6.74)**
		284 (4.78)	118 (4.40)	139 (4.35)	351 (5.08)	367 (6.45)
		359 (4.94)	77 (4.84)	111 (4.74)	324 (4.90)	388 (6.11)
BMI (n (mean))		N.A.	**218 (26.9)**	**298 (27.1)**	**387 (27.9)**	**332 (28.0)**
		280 (26.9)	111 (27.3)	139 (26.9)	348 (28.1)	357 (27.9)
		351 (27.3)	77 (28.2)	109 (27.0)	320 (27.4)	378 (28.0)
Any sleep problems		N.A.	**86 (38.4%)**	**146 (47.9%)**	**190 (48.6%)**	**210 (59.5%)**
		117 (40.5%)	50 (42.4%)	66 (47.5%)	175 (49.6%)	224 (59.9%)
		157 (43.0%)	28 (35.4%)	55 (48.7%)	163 (50.5%)	229 (58.3%)
Frequency of GP visit	Often/ Very Often	N.A.	**36 (16.1%)**	**72 (23.5%)**	**89 (22.9%)**	**105 (29.8%)**
		67 (23.1%)	16 (13.6%)	26 (18.4%)	84 (23.9%)	109 (29.4%)
		79 (21.8%)	12 (15.2%)	30 (26.6%)	65 (20.0%)	116 (29.6%)

Baseline factor		LA	HP	PGF	HPPGF	SA
SF-12 general health	Poor/ Fair	N.A.	**55 (24.7%)**	**93 (30.7%)**	**158 (40.8%)**	**221 (63.3%)**
		96 (66.3%)	29 (24.8%)	35 (25.2%)	126 (36.0%)	241 (65.0%)
		138 (38.6%)	20 (25.3%)	37 (33.0%)	100 (31.0%)	232 (59.5%)

Appendix F.3: Frequency (proportion (%'s)) of individuals in each hand state at baseline (bold), 3 years (underlined) and 6 years (standard font) for predictors in group three.

Baseline factor		LA	HP	PGF	HPPGF	SA
Previous hand injury	Yes	N.A.	**58 (26.5%)**	**95 (32.3%)**	**115 (30.2%)**	**117 (34.5%)**
		80 (28.9%)	30 (26.3%)	41 (29.7%)	109 (31.8%)	125 (34.6%)
		112 (31.9%)	23 (30.3%)	32 (28.8%)	93 (29.9%)	125 (32.6%)
Hand operation	Yes	N.A.	**33 (15.1%)**	**51 (17.4%)**	**67 (17.5%)**	**72 (21.1%)**
		49 (17.6%)	18 (15.8%)	24 (17.5%)	63 (18.2%)	69 (19.2%)
		60 (17.1%)	13 (16.9%)	18 (16.1%)	55 (17.5%)	77 (20.2%)
Excessive use	Yes	N.A.	**187 (83.5%)**	**244 (80.5%)**	**345 (88.0%)**	**305 (87.9%)**
		245 (84.8%)	97 (82.2%)	109 (77.9%)	310 (88.3%)	320 (87.0%)
		300 (83.3%)	66 (83.5%)	93 (82.3%)	280 (86.4%)	342 (87.7%)
Nodes	Yes	N.A.	**75 (34.9%)**	**159 (53.9%)**	**187 (49.9%)**	**185 (54.9%)**
		103 (37.7%)	46 (40.4%)	73 (53.7%)	171 (50.4%)	213 (59.2%)
		127 (37.0%)	31 (39.7%)	58 (53.7%)	161 (51.8%)	229 (60.0%)
Previous 12 month hand pain duration	≥3 months	N.A.	**139 (62.9%)**	**166 (57.2%)**	**293 (75.3%)**	**308 (87.5%)**
		174 (61.7%)	79 (67.5%)	88 (65.2%)	250 (71.8%)	315 (85.1%)
		226 (64.2%)	53 (68.0%)	67 (60.4%)	240 (74.5%)	320 (82.3%)

Baseline factor		LA	HP	PGF	HPPGF	SA
Bilateral hand pain	≥3 months	N.A.	**132 (60.8%)**	**197 (68.6%)**	**277 (71.0%)**	**279 (79.3%)**
		163 (58.4%)	76 (65.0%)	94 (70.7%)	255 (73.3%)	297 (80.5%)
		216 (61.7%)	44 (56.4%)	70 (64.2%)	243 (75.9%)	312 (80.2%)
Pain in two or more hand joints	All/ Most/ Some days	N.A.	**113 (51.6%)**	**132 (44.3%)**	**272 (71.0%)**	**301 (86.7%)**
		323 (23.1%)	89 (39.9%)	86 (44.1%)	273 (60.9%)	313 (77.9%)
		364 (26.1%)	69 (35.2%)	81 (45.0%)	249 (55.0%)	321 (73.1%)
Impact of hand problem	Fair/ Poor/ Very poorly	N.A.	**28 (12.6%)**	**58 (19.2%)**	**115 (29.4%)**	**206 (59.5%)**
		67 (23.3%)	21 (17.8%)	31 (22.3%)	94 (26.8%)	194 (52.9%)
		84 (23.3%)	18 (22.8%)	36 (32.1%)	83 (25.8%)	186 (47.9%)
RA		N.A.	**45 (20.1%)**	**41 (13.4%)**	**76 (19.3%)**	**65 (18.4%)**
		43 (14.8%)	29 (24.6%)	19 (13.5%)	61 (17.2%)	75 (20.0%)
		60 (16.4%)	13 (16.5%)	19 (16.8%)	58 (17.8%)	77 (19.5%)
Frequency of medication use	All/ Most days	N.A.	**22 (9.9%)**	**41 (13.8%)**	**95 (24.7%)**	**187 (54.1%)**
		43 (15.3%)	10 (8.6%)	20 (14.8%)	97 (27.9%)	175 (47.3%)
		67 (19.0%)	9 (11.5%)	22 (20.2%)	78 (24.2%)	169 (43.6%)

Appendix F.4: Frequency (proportion (%'s)) of individuals in each hand state at baseline (bold), 3 years (underlined) and 6 years (standard font) for predictors in group four.

Baseline factor	LA	HP	PGF	HPPGF	SA
Any comorbidity	N.A.	**117**	**161**	**237**	**239**
		(52.2%)	**(52.4%)**	**(60.2%)**	**(67.5%)**
	165	60	76	212	241
	(56.7%)	(50.9%)	(53.9%)	(59.9%)	(64.3%)
	207	38	61	196	252
	(56.6%)	(48.1%)	(54.0%)	(60.1%)	(63.8%)

Appendix F.5: Results from group one: Demographic predictors of latent state at 6 years, adjusting for baseline state only (RRR (95% CI)).

Baseline factor		LA	HP	PGF	HPPGF	SA
Gender	Male	1.00	1.75 (1.03,2.95)	0.89 (0.55,1.44)	0.72 (0.51,0.99)	0.56 (0.39,0.79)
Age (years)	50-64	1.00	1.00	1.00	1.00	1.00
	65-74	1.00	0.48 (0.26,0.88)	0.97 (0.61,1.55)	0.78 (0.56,1.09)	0.97 (0.69,1.37)
	75+	1.00	0.24 (0.07,0.80)	0.48 (0.22,1.04)	0.40 (0.23,0.70)	0.81 (0.50,1.32)
Lived alone	Yes	1.00	0.73 (0.36,1.48)	0.68 (0.39,1.20)	0.91 (0.62,1.32)	0.88 (0.60,1.28)
Employment Status	Retired	1.00	1.00	1.00	1.00	1.00
	Employed	1.00	2.53 (1.41,4.51)	1.31 (0.78,2.22)	1.60 (1.10,2.33)	0.92 (0.60,1.42)
	Other	1.00	1.72 (0.86,3.44)	0.84 (0.47,1.50)	1.34 (0.91,1.96)	1.06 (0.72,1.55)
Social Class	Higher managerial/ Professional	1.00	1.00	1.00	1.00	1.00
	Intermediate	1.00	0.81 (0.40,1.61)	0.50 (0.27,0.94)	0.72 (0.46,1.13)	0.63 (0.38,1.04)
	Routine/ Manual	1.00	0.64 (0.35,1.19)	0.56 (0.33,0.96)	0.64 (0.43,0.96)	0.82 (0.53,1.27)

Appendix F.6: Results from group two: Demographic predictors of latent state at 6 years, adjusting for baseline state only (RRR (95% CI)).

Baseline factor		LA	HP	PGF	HPPGF	SA
ACR Widespread pain		1.00	1.49 (0.90,2.48)	1.23 (0.79,1.91)	1.26 (0.93,1.72)	1.40 (1.01,1.93)
HADS	Depression	1.00	1.05 (0.97,1.13)	0.98 (0.91,1.04)	0.98 (0.94,1.03)	1.01 (0.97,1.06)
BMI		1.00	1.04 (1.00,1.09)	0.99 (0.94,1.03)	1.00 (0.97,1.03)	1.01 (0.99,1.04)
Any sleep problems		1.00	0.83 (0.49,1.40)	1.22 (0.79,1.89)	1.32 (0.97,1.79)	1.50 (1.10.2.07)
Frequency of GP visit	Occasionally/ Seldom/ Never	1.00	1.00	1.00	1.00	1.00
	Often/ Very Often	1.00	0.75 (0.38,1.49)	1.22 (0.74,2.01)	0.86 (0.59,1.25)	1.26 (0.88,1.81)
SF-12 general health	Good/ Very Good/ Excellent	1.00	1.00	1.00	1.00	1.00
	Poor/ Fair	1.00	0.69 (0.39,1.22)	0.79 (0.50,1.27)	0.66 (0.48,0.92)	1.46 (1.06,2.03)

Appendix F.7: Results from group three: Demographic predictors of latent state at 6 years, adjusting for baseline state only (RRR (95% CI)).

Baseline factor		LA	HP	PGF	HPPGF	SA
Previous hand injury	Yes	1.00	1.06 (0.61,1.84)	0.80 (0.50,1.30)	0.88 (0.63,1.23)	0.90 (0.64,1.27)
Hand operation	Yes	1.00	1.08 (0.55,2.12)	0.90 (0.50,1.63)	1.01 (0.67,1.52)	1.08 (0.71,1.62)
Excessive use	Yes	1.00	1.01 (0.51,1.99)	1.02 (0.57,1.82)	1.25 (0.81,1.91)	1.26 (0.80,1.99)
Nodes	Yes	1.00	1.36 (0.81,2.29)	1.70 (1.08,2.66)	1.72 (1.25,2.36)	2.36 (1.69,3.28)
Previous 12 month hand pain duration	<3 months	1.00	1.00	1.00	1.00	1.00
	≥3 months	1.00	1.34 (0.78,2.31)	1.00 (0.63,1.59)	1.64 (1.16,2.32)	1.61 (1.10,2.35)
Bilateral hand pain	Yes	1.00	0.89 (0.53,1.50)	1.10 (0.69,1.74)	1.93 (1.37,2.71)	2.11 (1.47,3.03)
Pain in two or more hand joints	No/ Few days	1.00	1.00	1.00	1.00	1.00
	All/ Most/ Some days	1.00	1.10 (0.65,1.85)	1.26 (0.79,2.01)	1.56 (1.12,2.17)	1.75 (1.22,2.53)
Impact of hand problem	Very well/ Well	1.00	1.00	1.00	1.00	1.00
	Fair/ Poor/ Very poorly	1.00	1.57 (0.84,2.93)	1.77 (1.07,2.92)	1.08 (0.75,1.57)	1.46 (1.02,2.09)
RA		1.00	0.87 (0.44,1.71)	1.21 (0.67,2.16)	1.14 (0.76,1.71)	1.30 (0.86,1.97)
Frequency of medication use	Some/ Few/ No days	1.00	1.00	1.00	1.00	1.00
	All/ Most days	1.00	0.82 (0.38,1.80)	1.22 (0.69,2.16)	1.35 (0.91,1.99)	1.61 (1.11,2.35)

Appendix F.8: Results from group four: Demographic predictors of latent state at 6 years, adjusting for baseline state only (RRR (95% CI)).

Baseline factor	LA	HP	PGF	HPPGF	SA
Any comorbidity	1.00	0.76 (0.46,1.26)	0.95 (0.61,1.47)	1.14 (0.84,1.56)	1.08 (0.79,1.50)

Appendix G: Trajectory univariable analysis

Appendix G.1: Results from group one: Demographic predictors of 6 year trajectory membership in the intermediate groups, adjusting for baseline state only (RRR (95% CI)).

Baseline factor		Improvers	Stable	Deterioraters	Fluctuaters
Gender	Male	1.00	0.90 (0.52,1.54)	0.42 (0.27,0.66)	0.74 (0.45,1.22)
Age (years)	50-64	1.00	1.00	1.00	1.00
	65-74	1.00	0.92 (0.51,1.67)	1.05 (0.66,1.68)	1.49 (0.87,2.56)
	75+	1.00	0.33 (0.09,1.16)	1.05 (0.52,2.10)	1.44 (0.66,3.18)
Lived alone	Yes	1.00	0.88 (0.42,1.84)	1.23 (0.71,2.12)	0.64 (0.31,1.32)
Employment Status	Retired	1.00	1.00	1.00	1.00
	Employed	1.00	1.27 (0.70,2.29)	0.82 (0.51,1.32)	0.83 (0.47,1.47)
	Other	1.00	0.87 (0.39,1.93)	0.94 (0.52,1.68)	1.05 (0.53,2.06)
Social Class	Higher managerial/ Professional	1.00	1.00	1.00	1.00
	Intermediate	1.00	0.70 (0.34,1.44)	0.86 (0.48,1.54)	0.64 (0.32,1.27)
	Routine/ Manual	1.00	0.79 (0.41,1.50)	1.20 (0.71,2.03)	1.02 (0.56,1.85)

Appendix G.2: Results from group two: General health predictors of 6 year trajectory membership in the intermediate groups, adjusting for baseline state only (RRR (95% CI)).

Baseline factor		Improvers	Stable	Deteriorators	Fluctuaters
ACR Widespread pain		1.00	1.30 (0.76,2.24)	1.57 (1.03,2.40)	1.49 (0.91,2.45)
HADS	Depression	1.00	1.00 (0.92,1.10)	1.07 (1.00,1.15)	1.06 (0.98,1.15)
BMI		1.00	1.02 (0.96,1.08)	0.99 (0.94,1.04)	1.06 (1.01,1.12)
Any sleep problems		1.00	0.98 (0.56,1.69)	1.84 (1.20,2.82)	0.94 (0.57,1.56)
Frequency of GP visit	Occasionally/ Seldom/ Never	1.00	1.00	1.00	1.00
	Often/ Very Often	1.00	0.65 (0.31,1.36)	1.00 (0.59,1.68)	1.37 (0.76,2.45)
SF-12 general health	Good/ Very Good/ Excellent	1.00	1.00	1.00	1.00
	Poor/ Fair	1.00	0.57 (0.29,1.11)	1.14 (0.72,1.81)	1.13 (0.66,1.94)

Appendix G.3: Results from group three: Hand-specific predictors of 6 year trajectory membership in the intermediate groups, adjusting for baseline state only (RRR (95% CI)).

Baseline factor		Improvers	Stable	Deteriorators	Fluctuaters
Previous hand injury	Yes	1.00	1.42 (0.79,2.56)	1.11 (0.69,1.79)	1.51 (0.87,2.59)
Hand operation	Yes	1.00	1.46 (0.71,2.99)	1.19 (0.66,2.16)	1.28 (0.65,2.55)
Excessive use	Yes	1.00	1.03 (0.52,2.03)	1.18 (0.69,2.03)	1.19 (0.63,2.28)
Nodes	Yes	1.00	1.88 (1.08,3.27)	2.09 (1.35,3.24)	1.20 (0.72,2.01)
Previous 12 month hand pain duration	<3 months	1.00	1.00	1.00	1.00
	≥3 months	1.00	1.34 (0.77,2.33)	1.79 (1.51,2.78)	1.10 (0.66,1.82)
Bilateral hand pain	Yes	1.00	1.18 (0.67,2.07)	2.11 (1.32,3.38)	0.96 (0.58,1.61)
Pain in two or more hand joints	No/ Few days	1.00	1.00	1.00	1.00
	All/ Most/ Some days	1.00	1.35 (0.78,2.33)	1.61 (1.05,2.47)	1.04 (0.62,1.72)
Impact of hand problem	Very well/ Well	1.00	1.00	1.00	1.00
	Fair/ Poor/ Very poorly	1.00	1.43 (0.70,2.91)	1.25 (0.70,2.23)	1.16 (0.58,2.30)
RA		1.00	0.52 (0.24,1.14)	0.72 (0.41,1.24)	0.69 (0.36,1.35)
Frequency of medication use	Some/ Few/ No days	1.00	1.00	1.00	1.00
	All/ Most day	1.00	1.10 (0.43,2.86)	2.03 (1.03,3.99)	1.62 (0.73,3.60)

Appendix G.4: Results from group four: Additional health predictors of 6 year trajectory membership in the intermediate groups, adjusting for baseline state only (RRR (95% CI)).

Baseline factor	Improvers	Stable	Deteriorators	Fluctuaters
Any comorbidity	1.00	0.91 (0.53,1.56)	1.52 (1.00,2.33)	0.88 (0.54,1.44)

Appendix G.5: Results from group one: Demographic predictors of 6 year trajectory membership in the worst groups, adjusting for baseline state only (RRR (95% CI)).

Baseline factor		Improvers	Stable	Fluctuaters
Gender	Male	1.00	0.63 (0.44,0.89)	0.94 (0.58,1.54)
Age (years)	50-64	1.00	1.00	1.00
	65-74	1.00	1.00 (0.70,1.42)	1.07 (0.64,1.78)
	75+	1.00	0.55 (0.32,0.94)	0.91 (0.44,1.87)
Lived alone	Yes	1.00	1.01 (0.69,1.47)	1.01 (0.58,1.75)
Employment Status	Retired	1.00	1.00	1.00
	Employed	1.00	1.01 (0.65,1.57)	0.94 (0.49,1.81)
	Other	1.00	1.32 (0.89,1.94)	1.21 (0.69,2.11)
Social Class	Higher managerial/ Professional	1.00	1.00	1.00
	Intermediate	1.00	0.76 (0.44,1.30)	0.82 (0.38,1.75)
	Routine/ Manual	1.00	0.64 (0.40,1.03)	0.64 (0.33,1.26)

Appendix G.6: Results from group two: General health predictors of 6 year trajectory membership in the worst groups, adjusting for baseline state only (RRR (95% CI)).

Baseline factor		Improvers	Stable	Fluctuaters
ACR Widespread pain		1.00	1.53 (1.09,2.14)	1.19 (0.74,1.91)
HADS	Depression	1.00	1.02 (0.97,1.06)	1.03 (0.96,1.09)
BMI		1.00	1.01 (0.98,1.04)	1.00 (0.95,1.04)
Any sleep problems		1.00	1.41 (1.02,1.96)	1.24 (0.77,1.98)
Frequency of GP visit	Occasionally/ Seldom/ Never	1.00	1.00	1.00
	Often/ Very Often	1.00	1.27 (0.86,1.86)	1.32 (0.77,2.24)
SF-12 general health	Good/ Very Good/ Excellent	1.00	1.00	1.00
	Poor/ Fair	1.00	1.39 (1.00,1.94)	1.08 (0.67,1.74)

Appendix G.7: Results from group three: Hand-specific predictors of 6 year trajectory membership in the worst groups, adjusting for baseline state only (RRR (95% CI)).

Baseline factor		Improvers	Stable	Fluctuaters
Previous hand injury	Yes	1.00	1.06 (0.74,1.51)	0.82 (0.48,1.39)
Hand operation	Yes	1.00	1.33 (0.87,2.03)	0.74 (0.38,1.47)
Excessive use	Yes	1.00	1.26 (0.78,2.05)	1.72 (0.79,3.77)
Nodes	Yes	1.00	2.25 (1.60,3.17)	1.54 (0.95,2.51)
Previous 12 month hand pain duration	<3 months	1.00	1.00	1.00
	≥3 months	1.00	1.89 (1.26,2.84)	1.33 (0.75,2.34)
Bilateral hand pain	Yes	1.00	2.43 (1.67,3.53)	1.09 (0.66,1.78)
Pain in two or more hand joints	No/ Few days	1.00	1.00	1.00
	All/ Most/ Some days	1.00	2.17 (1.47,3.21)	1.22 (0.71,2.08)
Impact of hand problem	Very well/ Well	1.00	1.00	1.00
	Fair/ Poor/ Very poorly	1.00	1.47 (1.05,2.05)	1.04 (0.64,1.69)
RA		1.00	1.48 (0.96,2.29)	1.09 (0.58,2.05)
Frequency of medication use	Some/ Few/ No days	1.00	1.00	1.00
	All/ Most day	1.00	1.99 (1.40,2.83)	0.91 (0.54,1.53)

Appendix G.8: Results from group four: Additional health predictors of 6 year trajectory membership in the worst groups, adjusting for baseline state only (RRR (95% CI)).

Baseline factor	Improvers	Stable	Fluctuaters
Any comorbidity	1.00	1.16 (0.82,1.62)	0.92 (0.57,1.50)

Appendix H: Mental health state development tables

Appendix H.1: Model fit parameters with 14 HADS indicator variables for the first removal stage.

Number of latent states	BIC	Entropy	Lowest sample size
2	130,667	0.918	21.8%
3	125,964	0.848	9.5%
4	**123,224**	**0.881**	**7.9%**
5	122,313	0.853	6.8%
6	121,479	0.851	5.3%

Appendix H.2: Model fit parameters when removing one indicator in turn from model with 14 indicators and four states.

Removing the term:	BIC	Entropy
Feel tense or 'wound up'	114,804	0.881
Still enjoy the things I used to enjoy	115,819	0.870
Frightened feeling as if something awful is about to happen	112,922	0.875
Can laugh and see the funny side of things	118,902	0.880
Worrying thoughts go through my mind	112,344	0.875
Feel cheerful	119,914	0.879
Can sit at ease and feel relaxed	115,112	0.879
Feel as if I am slowed down	107,599	0.887
Frightened feeling like 'butterflies' in my stomach	117,688	0.877
Lost interest in my appearance	116,047	0.878
Feel restless as if I have to be on the move	**105,954**	**0.883**
Look forward with enjoyment to things	115,446	0.875
Get sudden feelings of panic	116,725	0.875
Can enjoy a good book or radio or television programme	119,452	0.879

Appendix H.3: Model fit parameters with 13 HADS indicator variables for the second removal stage.

Number of latent states	BIC	Entropy	Lowest sample size
2	113,174	0.917	22.1%
3	108,734	0.852	8.9%
4	**105,954**	**0.883**	**7.6%**
5	105,119	0.871	5.0%
6	104,593	0.863	2.6%

Appendix H.4: Model fit parameters when removing one indicator in turn from model with 13 indicators and four states.

Removing the term:	BIC	Entropy
Feel tense or 'wound up'	97,448	0.884
Still enjoy the things I used to enjoy	98,593	0.873
Frightened feeling as if something awful is about to happen	95,766	0.879
Can laugh and see the funny side of things	101,661	0.882
Worrying thoughts go through my mind	95,047	0.878
Feel cheerful	102,671	0.882
Can sit at ease and feel relaxed	97,704	0.882
Feel as if I am slowed down	**90,380**	**0.891**
Frightened feeling like 'butterflies' in my stomach	100,436	0.879
Lost interest in my appearance	98,775	0.881
Look forward with enjoyment to things	98,210	0.878
Get sudden feelings of panic	99,423	0.878
Can enjoy a good book or radio or television programme	102,173	0.882

Appendix H.5: Model fit parameters with 12 HADS indicator variables for the third removal stage.

Number of latent states	BIC	Entropy	Lowest sample size
2	96,032	0.919	20.0%
3	92,482	0.864	6.9%
4	**90,380**	**0.891**	**6.0%**
5	89,798	0.874	3.3%
6	89,607	0.857	2.2%

Appendix H.6: Model fit parameters when removing one indicator in turn from model with 12 indicators and four states.

Removing the term:	BIC	Entropy
Feel tense or 'wound up'	81,909	0.891
Still enjoy the things I used to enjoy	82,469	0.885
Frightened feeling as if something awful is about to happen	80,373	0.882
Can laugh and see the funny side of things	86,164	0.887
Worrying thoughts go through my mind	**79,620**	**0.883**
Feel cheerful	87,036	0.890
Can sit at ease and feel relaxed	82,253	0.891
Frightened feeling like 'butterflies' in my stomach	85,013	0.885
Lost interest in my appearance	83,187	0.890
Look forward with enjoyment to things	82,604	0.886
Get sudden feelings of panic	84,015	0.883
Can enjoy a good book or radio or television programme	86,627	0.891

Appendix H.7: Model fit parameters with 11 HADS indicator variables for the fourth removal stage.

Number of latent states	BIC	Entropy	Lowest sample size
2	83,941	0.915	18.0%
3	81,350	0.847	6.9%
4	**79,620**	**0.883**	**4.3%**
5	79,247	0.858	3.1%
6	79,126	0.846	2.1%

Appendix H.8: Model fit parameters when removing one indicator in turn from model with 11 indicators and four states.

Removing the term:	BIC	Entropy
Feel tense or 'wound up'	70,693	0.883
Still enjoy the things I used to enjoy	71,775	0.879
Frightened feeling as if something awful is about to happen	**68,707**	**0.874**
Can laugh and see the funny side of things	75,469	0.880
Feel cheerful	76,335	0.884
Can sit at ease and feel relaxed	71,332	0.890
Frightened feeling like 'butterflies' in my stomach	74,228	0.873
Lost interest in my appearance	72,456	0.885
Look forward with enjoyment to things	71,907	0.881
Get sudden feelings of panic	73,079	0.870
Can enjoy a good book or radio or television programme	75,406	0.886

Appendix H.9: Model fit parameters with 10 HADS indicator variables for the fifth removal stage.

Number of latent states	BIC	Entropy	Lowest sample size
2	72,111	0.909	17.0%
3	69,947	0.851	5.9%
4	**68,707**	**0.874**	**4.4%**
5	68,520	0.856	3.7%
6	68,415	0.881	3.2%

Appendix H.10: Model fit parameters when removing one indicator in turn from model with 10 indicators and four states.

Removing the term:	BIC	Entropy
Feel tense or 'wound up'	59,556	0.879
Still enjoy the things I used to enjoy	60,928	0.871
Can laugh and see the funny side of things	64,646	0.875
Feel cheerful	65,544	0.880
Can sit at ease and feel relaxed	**60,444**	**0.878**
Frightened feeling like 'butterflies' in my stomach	63,065	0.871
Lost interest in my appearance	61,552	0.871
Look forward with enjoyment to things	61,066	0.876
Get sudden feelings of panic	61,573	0.874
Can enjoy a good book or radio or television programme	65,034	0.877

Appendix H.11: Model fit parameters with nine HADS indicator variables for the sixth removal stage.

Number of latent states	BIC	Entropy	Lowest sample size
2	63,441	0.899	17.4%
3	61,853	0.854	4.0%
4	**60,444**	**0.878**	**3.5%**
5	60,375	0.852	3.0%
6	60,403	0.842	1.8%

Appendix H.12: Model fit parameters when removing one indicator in turn from model with nine indicators and four states.

Removing the term:	BIC	Entropy
Feel tense or 'wound up'	50,790	0.881
Still enjoy the things I used to enjoy	52,660	0.877
Can laugh and see the funny side of things	56,383	0.874
Feel cheerful	57,159	0.882
Frightened feeling like 'butterflies' in my stomach	54,927	0.865
Lost interest in my appearance	53,303	0.879
Look forward with enjoyment to things	52,863	0.873
Get sudden feelings of panic	53,533	0.865
Can enjoy a good book or radio or television programme	56,712	0.878

Appendix H.13: Item-response probabilities of model with nine indicators, four states and removing 'feel tense or 'wound up' '.

Removing the term: Feel tense or 'wound up'	State 1	State 2	State 3	State 4
Baseline state proportion	0.740	0.133	0.093	0.035
Item-response probabilities				
Still enjoy the things I used to enjoy	0.013	0.372	0.113	0.812
Can laugh and see the funny side of things	0.004	0.129	0.078	0.646
Feel cheerful	0.002	0.069	0.062	0.554
Frightened feeling like 'butterflies' in my stomach	0.009	0.043	0.588	0.659
Lost interest in my appearance	0.026	0.244	0.117	0.513
Look forward with enjoyment to things	0.016	0.398	0.203	0.908
Get sudden feelings of panic	0.014	0.071	0.802	0.696
Can enjoy a good book or radio or television programme	0.008	0.070	0.052	0.285

Appendix H.14: Item-response probabilities of model with nine indicators, four states and removing 'still enjoy the things I used to enjoy'.

Removing the term: Still enjoy the things I used to enjoy	State 1	State 2	State 3	State 4
Baseline state proportions	0.732	0.146	0.089	0.032
Item-response probabilities				
Feel tense or 'wound up'	0.054	0.390	0.674	0.957
Can laugh and see the funny side of things	0.002	0.153	0.060	0.630
Feel cheerful	0.001	0.088	0.044	0.558
Frightened feeling like 'butterflies' in my stomach	0.010	0.047	0.582	0.744
Lost interest in my appearance	0.022	0.269	0.092	0.514
Look forward with enjoyment to things	0.014	0.410	0.188	0.859
Get sudden feelings of panic	0.011	0.101	0.773	0.804
Can enjoy a good book or radio or television programme	0.005	0.091	0.042	0.286

Appendix H.15: Item-response probabilities of model with nine indicators, four states and removing "Look forward to things with enjoyment to things".

Removing the term: Look forward with enjoyment to things	State 1	State 2	State 3	State 4
Baseline state proportions	0.722	0.155	0.091	0.031
Item-response probabilities				
Feel tense or 'wound up'	0.057	0.352	0.681	0.963
Still enjoy the things I used to enjoy	0.013	0.349	0.106	0.779
Can laugh and see the funny side of things	0.003	0.140	0.066	0.637
Feel cheerful	0.001	0.082	0.055	0.545
Frightened feeling like 'butterflies' in my stomach	0.009	0.050	0.574	0.746
Lost interest in my appearance	0.023	0.242	0.113	0.502
Get sudden feelings of panic	0.012	0.082	0.788	0.789
Can enjoy a good book or radio or television programme	0.005	0.089	0.039	0.303

Appendix H.16: Model parameters when including one indicator to model with nine indicators and four states, previously removed, in turn.

Including the term:	BIC	Entropy
Frightened feeling as if something awful is about to happen	71,332	0.890
Worrying thoughts go through my mind	72,241	0.885
Can sit at ease and feel relaxed	68,707	0.874
Feel as if I am slowed down	75,718	0.874
Feel restless as if I have to be on the move	77,920	0.846

Appendix H.17: Item-response probabilities for each indicator at all three time points split by latent state in the mental health model.

Item-response probabilities		No anxiety/ depression	Mild depression	High anxiety	Anxiety & depression
Feel tense or 'wound up'	Time 1	0.060	0.331	0.673	0.942
	Time 2	0.033	0.222	0.517	0.860
	Time 3	0.019	0.179	0.535	0.844
Still enjoy the things I used to enjoy	Time 1	0.015	0.381	0.097	0.801
	Time 2	0.011	0.361	0.089	0.746
	Time 3	0.016	0.495	0.097	0.749
Can laugh and see the funny side of things	Time 1	0.003	0.146	0.070	0.605
	Time 2	0.004	0.100	0.031	0.458
	Time 3	0.002	0.116	0.038	0.513
Feel cheerful	Time 1	0.001	0.082	0.056	0.522
	Time 2	0.001	0.051	0.041	0.476
	Time 3	0.001	0.076	0.066	0.518
Frightened feeling like 'butterflies' in my stomach	Time 1	0.008	0.043	0.533	0.688
	Time 2	0.007	0.022	0.398	0.591
	Time 3	0.005	0.031	0.370	0.572
Lost interest in my appearance	Time 1	0.027	0.252	0.111	0.483
	Time 2	0.018	0.188	0.101	0.406
	Time 3	0.018	0.226	0.088	0.468
Look forward with enjoyment to things	Time 1	0.014	0.421	0.188	0.889
	Time 2	0.009	0.336	0.104	0.832
	Time 3	0.010	0.454	0.129	0.853
Get sudden feelings of panic	Time 1	0.012	0.073	0.734	0.736
	Time 2	0.009	0.045	0.620	0.686
	Time 3	0.007	0.029	0.627	0.727
Can enjoy a good book or radio or television programme	Time 1	0.007	0.077	0.052	0.273
	Time 2	0.006	0.068	0.018	0.277
	Time 3	0.009	0.087	0.036	0.223

Appendix I: The Hospital Anxiety and Depression Scale (HADS)

How you have felt in the past week

The next questions are about how you feel at the moment. Please read each item and put a cross next to the reply that comes closest to how you have been feeling in **the past week**. Don't take too long over your replies; your immediate reaction to each item will usually be more accurate than a long thought out response.

1. I feel tense or 'wound up':

Most of the time	A lot of the time	From time to time, occasionally	Not at all
☐	☐	☐	☐

2. I still enjoy the things I used to enjoy:

Definitely as much	Not quite as much	Only a little	Hardly at all
☐	☐	☐	☐

3. I get a sort of frightened feeling as if something awful is about to happen:

Very definitely and quite badly	Yes, but not too badly	A little, but it doesn't worry me	Not at all
☐	☐	☐	☐

4. I can laugh and see the funny side of things:

As much as I always could	Not quite so much now	Definitely not so much now	Not at all
☐	☐	☐	☐

5. Worrying thoughts go through my mind:

A great deal of the time	A lot of the time	Not too often	Very little
☐	☐	☐	☐

6. I feel cheerful:

Never	Not often	Sometimes	Most of the time
☐	☐	☐	☐

7. I can sit at ease and feel relaxed:

Definitely	Usually	Not often	Not at all
☐	☐	☐	☐

8. I feel as if I am slowed down:

Nearly all the time	Very often	Sometimes	Not at all
☐	☐	☐	☐

9. I get a sort of frightened feeling like ‘butterflies’ in my stomach:

Not at all	Occasionally	Quite often	Very often
☐	☐	☐	☐

10. I have lost interest in my appearance:

Definitely	I don’t take as much care as I should	I may not take quite as much care	I take just as much care as ever
☐	☐	☐	☐

11. I feel restless as if I have to be on the move:

Very much indeed	Quite a lot	Not very much	Not at all
☐	☐	☐	☐

12. I look forward with enjoyment to things:

As much as I ever did	Rather less than I used to	Definitely less than I used to	Hardly at all
☐	☐	☐	☐

13. I get sudden feelings of panic:

Very often indeed	Quite often	Not very often	Not at all
☐	☐	☐	☐

14. I can enjoy a good book or radio or television programme:

Often	Sometimes	Not often	Very seldom
☐	☐	☐	☐

Appendix J: ALTA tables

Tables Legend:

T1/ T2/ T3 = Time 1/ Time 2/ Time 3

MH = Mental health state

hand= Hand state

Appendix J.1: β estimates; Cross-sectional time n mental health state dependent on time n hand state.

Baseline (Time 1):

	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.8519	0.0497	0.0482	0.0503
High pain	0.1696	0.7138	0.0385	0.0781
Poor gross function	0.1065	0.1322	0.7191	0.0423
High pain & poor gross function	0.1686	0.1719	0.1688	0.4908
Severely affected	0.1569	0.1540	0.2728	0.4163

3 years (Time 2):

	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.4506	0.1655	0.1486	0.2354
High pain	0.3401	0.3430	0.1719	0.1450
Poor gross function	0.2488	0.3171	0.3105	0.1235
High pain & poor gross function	0.1211	0.3284	0.2565	0.2941
Severely affected	0.2693	0.1771	0.1313	0.4224

6 years (Time 3):

	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.4678	0.1166	0.1559	0.2598
High pain	0.3319	0.4830	0.0902	0.0949
Poor gross function	0.2577	0.3209	0.3334	0.0880
High pain & poor gross function	0.1522	0.2813	0.2618	0.3047
Severely affected	0.2007	0.0978	0.3465	0.3550

Appendix J.2: ε estimates; Time n hand state dependent on time n-1 mental health and time n-1 hand state (bold cells represent stability).

Time 1 to Time 2

In Least affected Time 1:

	Least affected	High pain	Poor gross function	High pain and poor gross function	Severely affected
No anx/ dep	**0.9347**	0.0387	0.0186	0.0067	0.0014
Mild dep	**0.8753**	0.0773	0.0126	0.0289	0.0059
High anx	**0.7853**	0.1545	0.0436	0.0166	0.0000
Anx & dep	**0.6924**	0.2127	0.0543	0.0348	0.0058

In High pain Time 1:

	Least affected	High pain	Poor gross function	High pain and poor gross function	Severely affected
No anx/ dep	0.3981	**0.3324**	0.1359	0.1066	0.0270
Mild dep	0.3899	**0.4360**	0.0798	0.0858	0.0085
High anx	0.5097	**0.2919**	0.1059	0.0476	0.0449
Anx & dep	0.4908	**0.1366**	0.2535	0.0434	0.0757

In Poor gross function Time 1:

	Least affected	High pain	Poor gross function	High pain and poor gross function	Severely affected
No anx/ dep	0.1636	0.3394	**0.3123**	0.1567	0.0279
Mild dep	0.1258	0.3914	**0.4387**	0.0442	0.0000
High anx	0.3579	0.0272	**0.4578**	0.1523	0.0048
Anx & dep	0.3673	0.2300	**0.2049**	0.1978	0.0000

In High pain and poor gross function Time 1:

	Least affected	High pain	Poor gross function	High pain and poor gross function	Severely affected
No anx/ dep	0.2207	0.0212	0.2185	**0.4993**	0.0403
Mild dep	0.2893	0.1165	0.1539	**0.3784**	0.0618
High anx	0.1017	0.1602	0.1655	**0.4861**	0.0865
Anx & dep	0.1110	0.1884	0.1541	**0.4275**	0.1190

In Severely affected Time 1:

	Least affected	High pain	Poor gross function	High pain and poor gross function	Severely affected
No anx/ dep	0.0000	0.0000	0.1944	0.3997	**0.4059**
Mild dep	0.1684	0.2356	0.1936	0.1158	**0.2866**
High anx	0.0362	0.1234	0.2736	0.1740	**0.3929**
Anx & dep	0.0233	0.0597	0.2803	0.1221	**0.5146**

Time 2 to Time 3

In Least affected Time 2:

	Least affected	High pain	Poor gross function	High pain and poor gross function	Severely affected
No anx/ dep	**0.9101**	0.0521	0.0252	0.0109	0.0017
Mild dep	**0.8365**	0.0954	0.0503	0.0178	0.0000
High anx	**0.8537**	0.0828	0.0230	0.0405	0.0000
Anx & dep	**0.7258**	0.1727	0.0636	0.0307	0.0072

In High pain Time 2:

	Least affected	High pain	Poor gross function	High pain and poor gross function	Severely affected
No anx/ dep	0.2759	**0.5556**	0.1070	0.0545	0.0070
Mild dep	0.2560	**0.6437**	0.0434	0.0456	0.0113
High anx	0.2634	**0.5692**	0.0774	0.0374	0.0525
Anx & dep	0.3049	**0.3329**	0.2163	0.0663	0.0796

In Poor gross function Time 2:

	Least affected	High pain	Poor gross function	High pain and poor gross function	Severely affected
No anx/ dep	0.1704	0.2926	**0.3399**	0.0968	0.1003
Mild dep	0.0673	0.3535	**0.3780**	0.0544	0.1468
High anx	0.3461	0.0530	**0.4796**	0.0921	0.0292
Anx & dep	0.3620	0.0000	**0.4533**	0.1847	0.0000

In High pain and poor gross function Time 2:

	Least affected	High pain	Poor gross function	High pain and poor gross function	Severely affected
No anx/ dep	0.2233	0.1628	0.2812	**0.2512**	0.0816
Mild dep	0.2807	0.0517	0.1525	**0.3968**	0.1183
High anx	0.0144	0.1389	0.2077	**0.4470**	0.1920
Anx & dep	0.1033	0.2080	0.0943	**0.4623**	0.1321

In Severely affected Time 2:

	Least affected	High pain	Poor gross function	High pain and poor gross function	Severely affected
No anx/ dep	0.1355	0.2788	0.3151	0.1805	**0.0901**
Mild dep	0.0872	0.0000	0.2528	0.4230	**0.2370**
High anx	0.0623	0.1423	0.3152	0.0000	**0.4803**
Anx & dep	0.0598	0.0000	0.2521	0.2777	**0.4104**

Appendix J.3: η estimates; Time n mental health state dependent on time n hand state, time n-1 mental health and time n-1 hand state (bold cells represent stability).

Time 1 to 2:

In Time 1 Least affected, Time 1 No anx/ dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	**0.9093**	0.0438	0.0171	0.0298
High pain	0.1804	0.6749	0.0591	0.0855
Poor gross function	0.0216	0.1168	0.7781	0.0835
High pain & poor gross function	0.1547	0.2900	0.1787	0.3766
Severely affected	0.5230	0.2491	0.0000	0.2280

In Time 1 Least affected, Time 1 Mild dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.4377	**0.3656**	0.0409	0.1557
High pain	0.2211	0.7789	0.0000	0.0000
Poor gross function	0.0000	0.0000	0.5049	0.4951
High pain & poor gross function	0.4199	0.1942	0.1812	0.2047
Severely affected	1.0000	0.0000	0.0000	0.0000

In Time 1 Least affected, Time 1 High anx:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.2479	0.0500	**0.4021**	0.3000
High pain	0.7098	0.1513	0.0000	0.1389
Poor gross function	0.3288	0.3423	0.2367	0.0922
High pain & poor gross function	0.0000	1.0000	0.0000	0.0000
Severely affected	0.0000	1.0000	0.0000	0.0000

In Time 1 Least affected, Time 1 Anx & dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.1660	0.0884	0.1376	**0.6080**
High pain	0.7793	0.1774	0.0176	0.0256
Poor gross function	0.1221	0.5450	0.2284	0.1045
High pain & poor gross function	0.0000	0.0000	0.6592	0.3408
Severely affected	0.0000	0.0000	0.0000	1.0000

In Time 1 High pain, Time 1 No anx/ dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.5434	0.0235	0.0391	0.3941
High pain	**0.8834**	0.0203	0.0325	0.0637
Poor gross function	0.3527	0.6473	0.0000	0.0000
High pain & poor gross function	0.0000	0.1914	0.7134	0.0952
Severely affected	0.0000	0.0000	0.0000	1.0000

In Time 1 High pain, Time 1 Mild dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.8453	0.0941	0.0482	0.0124
High pain	0.0153	**0.8871**	0.0562	0.0414
Poor gross function	0.3469	0.2540	0.3006	0.0985
High pain & poor gross function	0.0000	0.0525	0.0351	0.9124
Severely affected	0.2950	0.3848	0.3202	0.0000

In Time 1 High pain, Time 1 High anx:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.2689	0.6303	0.1008	0.0000
High pain	0.0678	0.1875	**0.5700**	0.1747
Poor gross function	0.4908	0.0000	0.5092	0.0000
High pain & poor gross function	0.0000	1.0000	0.0000	0.0000
Severely affected	1.0000	0.0000	0.0000	0.0000

In Time 1 High pain, Time 1 Anx & dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.6261	0.0171	0.1713	0.1855
High pain	0.0000	0.6335	0.3665	**0.0000**
Poor gross function	0.3806	0.4221	0.1973	0.0000
High pain & poor gross function	0.4937	0.5063	0.0000	0.0000
Severely affected	0.0000	0.5766	0.1355	0.2879

In Time 1 Poor gross function, Time 1 No anx/ dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.0000	0.2293	0.0000	0.7707
High pain	0.3205	0.6100	0.0695	0.0000
Poor gross function	**0.6843**	0.1416	0.1194	0.0547
High pain & poor gross function	0.1931	0.4571	0.1316	0.2183
Severely affected	0.0000	0.0000	0.0000	1.0000

In Time 1 Poor gross function, Time 1 Mild dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.4099	0.0899	0.1082	0.3919
High pain	0.6163	0.1715	0.0000	0.2122
Poor gross function	0.0000	**0.9678**	0.0322	0.0000
High pain & poor gross function	0.0000	0.2881	0.7119	0.0000
Severely affected	0.0000	0.0000	0.0000	1.0000

In Time 1 Poor gross function, Time 1 High anx:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.8435	0.0910	0.0393	0.0261
High pain	0.2733	0.5423	0.0863	0.0981
Poor gross function	0.0288	0.0176	**0.9091**	0.0445
High pain & poor gross function	0.1654	0.3378	0.1187	0.3782
Severely affected	0.7490	0.2510	0.0000	0.0000

In Time 1 Poor gross function, Time 1 Anx & dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.4569	0.3213	0.0000	0.2217
High pain	0.2028	0.0587	0.7385	0.0000
Poor gross function	0.0000	0.0000	0.5358	**0.4642**
High pain & poor gross function	0.3702	0.1992	0.4306	0.0000
Severely affected	1.0000	0.0000	0.0000	0.0000

In Time 1 High pain & poor gross function, Time 1 No anx/ dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.5145	0.0000	0.1288	0.3568
High pain	0.0000	0.0000	0.0000	1.0000
Poor gross function	0.3177	0.1020	0.2812	0.2992
High pain & poor gross function	**0.2444**	0.4315	0.2313	0.0929
Severely affected	0.0000	0.4564	0.0000	0.5436

In Time 1 High pain & poor gross function, Time 1 Mild dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.2462	0.0000	0.1631	0.5907
High pain	0.7127	0.2873	0.0000	0.0000
Poor gross function	0.0000	0.5636	0.1881	0.2483
High pain & poor gross function	0.1129	**0.6235**	0.1592	0.1044
Severely affected	0.0000	0.0000	1.0000	0.0000

In Time 1 High pain & poor gross function, Time 1 High anx:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.3679	0.2271	0.1883	0.2166
High pain	0.8738	0.0000	0.0000	0.1262
Poor gross function	0.2494	0.5433	0.2073	0.0000
High pain & poor gross function	0.0921	0.1053	**0.7921**	0.0105
Severely affected	0.0000	0.0000	0.0000	1.0000

In Time 1 High pain & poor gross function, Time 1 Anx & dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.8926	0.0387	0.0000	0.0687
High pain	0.0000	0.9581	0.0419	0.0000
Poor gross function	0.0000	0.1954	0.8046	0.0000
High pain & poor gross function	0.0171	0.0839	0.0000	**0.8990**
Severely affected	0.2227	0.2582	0.1472	0.3719

In Time 1 Severely affected, Time 1 No anx/ dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.0000	1.0000	0.0000	0.0000
High pain	0.0000	0.0000	1.0000	0.0000
Poor gross function	0.3257	0.3099	0.0000	0.3645
High pain & poor gross function	0.1577	0.0000	0.0000	0.8423
Severely affected	**0.2040**	0.0000	0.2280	0.5680

In Time 1 Severely affected, Time 1 Mild dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.2349	0.0000	0.3862	0.3789
High pain	0.5047	0.2639	0.0000	0.2314
Poor gross function	0.7299	0.2701	0.0000	0.0000
High pain & poor gross function	0.0000	0.4209	0.5791	0.0000
Severely affected	0.2530	**0.2258**	0.2969	0.2243

In Time 1 Severely affected, Time 1 High anx:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	1.0000	0.0000	0.0000	0.0000
High pain	0.2985	0.0000	0.0000	0.7015
Poor gross function	0.5003	0.0000	0.3779	0.1218
High pain & poor gross function	0.0000	0.0000	0.0000	1.0000
Severely affected	0.0924	0.0926	**0.3976**	0.4174

In Time 1 Severely affected, Time 1 Anx & dep:

T2 MH, on T2 hand, T1 MH, T1 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.0000	0.0000	1.0000	0.0000
High pain	0.1429	0.4575	0.3996	0.0000
Poor gross function	0.0959	0.9041	0.0000	0.0000
High pain & poor gross function	0.0000	0.3862	0.2079	0.4060
Severely affected	0.0465	0.0472	0.0998	**0.8065**

Time 2 to 3:

In Time 2 Least affected, Time 2 No anx/ dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	**0.9104**	0.0508	0.0137	0.0252
High pain	0.2208	0.6934	0.0438	0.0420
Poor gross function	0.2132	0.0491	0.6687	0.0690
High pain & poor gross function	0.0754	0.3015	0.0742	0.5489
Severely affected	0.4380	0.0000	0.0000	0.5620

In Time 2 Least affected, Time 2 Mild dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.4748	**0.2840**	0.0577	0.1835
High pain	0.5805	0.3736	0.0000	0.0459
Poor gross function	0.5017	0.0977	0.1577	0.2429
High pain & poor gross function	0.2459	0.4241	0.0000	0.3300
Severely affected	0.0000	0.0000	1.0000	0.0000

In Time 2 Least affected, Time 2 High anx:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.3056	0.0266	**0.4176**	0.2502
High pain	0.3305	0.3777	0.0000	0.2918
Poor gross function	0.8478	0.1522	0.0000	0.0000
High pain & poor gross function	0.4972	0.1602	0.1576	0.1850
Severely affected	0.0000	0.0000	1.0000	0.0000

In Time 2 Least affected, Time 2 Anx & dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.2550	0.1428	0.1003	**0.5019**
High pain	0.7694	0.1142	0.0950	0.0215
Poor gross function	0.4802	0.3820	0.1378	0.0000
High pain & poor gross function	0.2658	0.2764	0.4578	0.0000
Severely affected	0.5181	0.0000	0.4819	0.0000

In Time 2 High pain, Time 2 No anx/ dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.4939	0.0183	0.1285	0.3594
High pain	**0.9707**	0.0293	0.0000	0.0000
Poor gross function	0.1191	0.6739	0.2070	0.0000
High pain & poor gross function	0.0000	0.0000	1.0000	0.0000
Severely affected	1.0000	0.0000	0.0000	0.0000

In Time 2 High pain, Time 2 Mild dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.8633	0.0312	0.0406	0.0649
High pain	0.0086	**0.9138**	0.0347	0.0429
Poor gross function	0.3451	0.6549	0.0000	0.0000
High pain & poor gross function	0.0977	0.0629	0.0000	0.8393
Severely affected	0.0000	0.7270	0.0000	0.2730

In Time 2 High pain, Time 2 High anx:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.3938	0.3570	0.0566	0.1926
High pain	0.1782	0.6500	**0.1719**	0.0000
Poor gross function	0.0000	0.1944	0.4255	0.3800
High pain & poor gross function	0.0000	0.9125	0.0875	0.0000
Severely affected	0.4098	0.0000	0.0000	0.5902

In Time 2 High pain, Time 2 Anx & dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.3182	0.0965	0.2274	0.3580
High pain	0.2344	0.4491	0.0000	**0.3165**
Poor gross function	0.7472	0.2528	0.0000	0.0000
High pain & poor gross function	0.0000	0.4532	0.5468	0.0000
Severely affected	0.0000	0.0000	1.0000	0.0000

In Time 2 Poor gross function, Time 2 No anx/ dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.0000	0.5553	0.0858	0.3589
High pain	0.2210	0.4185	0.2241	0.1364
Poor gross function	**0.2742**	0.6609	0.0133	0.0517
High pain & poor gross function	0.4162	0.1856	0.3982	0.0000
Severely affected	0.0000	0.0000	0.9279	0.0721

In Time 2 Poor gross function, Time 2 Mild dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.4942	0.0000	0.1581	0.3477
High pain	0.3244	0.5691	0.0196	0.0869
Poor gross function	0.1450	**0.8006**	0.0545	0.0000
High pain & poor gross function	0.3652	0.1781	0.4567	0.0000
Severely affected	0.0000	0.0000	0.3452	0.6548

In Time 2 Poor gross function, Time 2 High anx:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.8795	0.1205	0.0000	0.0000
High pain	0.3285	0.5719	0.0996	0.0000
Poor gross function	0.0069	0.0290	**0.9580**	0.0061
High pain & poor gross function	0.1994	0.1781	0.1213	0.5012
Severely affected	0.0662	0.1198	0.4098	0.4042

In Time 2 Poor gross function, Time 2 Anx & dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.5606	0.3464	0.0000	0.0930
High pain	0.0000	0.0000	1.0000	0.0000
Poor gross function	0.0000	0.0000	0.6771	**0.3229**
High pain & poor gross function	0.0000	0.4609	0.1583	0.3808
Severely affected	0.0000	0.0000	0.0000	1.0000

In Time 2 High pain & poor gross function, Time 2 No anx/ dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.3972	0.1230	0.3244	0.1555
High pain	0.6013	0.0000	0.0000	0.3987
Poor gross function	0.4584	0.0000	0.5416	0.0000
High pain & poor gross function	**0.3350**	0.1297	0.3513	0.1840
Severely affected	0.3871	0.0000	0.3113	0.3016

In Time 2 High pain & poor gross function, Time 2 Mild dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.4455	0.0589	0.1334	0.3621
High pain	0.5270	0.0000	0.0000	0.4730
Poor gross function	0.0000	0.1209	0.7962	0.0829
High pain & poor gross function	0.0000	**0.4743**	0.4542	0.0714
Severely affected	0.0772	0.1083	0.2473	0.5672

In Time 2 High pain & poor gross function, Time 2 High anx:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	1.0000	0.0000	0.0000	0.0000
High pain	0.8783	0.1217	0.0000	0.0000
Poor gross function	0.0839	0.6332	0.2829	0.0000
High pain & poor gross function	0.0000	0.3254	**0.6746**	0.0000
Severely affected	0.0000	0.0000	0.0000	1.0000

In Time 2 High pain & poor gross function, Time 2 Anx & dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.7626	0.1203	0.0000	0.1171
High pain	0.0000	0.8422	0.1158	0.0420
Poor gross function	0.0000	0.2072	0.7928	0.0000
High pain & poor gross function	0.0320	0.0163	0.0315	**0.9202**
Severely affected	0.1177	0.0000	0.4337	0.4485

In Time 2 Severely affected, Time 2 No anx/ dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.4907	0.0000	0.0000	0.5093
High pain	0.4637	0.5363	0.0000	0.0000
Poor gross function	0.2228	0.2181	0.3414	0.2177
High pain & poor gross function	0.0000	0.0000	0.0000	1.0000
Severely affected	**1.0000**	0.0000	0.0000	0.0000

In Time 2 Severely affected, Time 2 Mild dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.0000	0.0000	0.0000	1.0000
High pain	0.0000	1.0000	0.0000	0.0000
Poor gross function	0.6284	0.3716	0.0000	0.0000
High pain & poor gross function	0.5134	0.0000	0.0000	0.4866
Severely affected	0.0000	**1.0000**	0.0000	0.0000

In Time 2 Severely affected, Time 2 High anx:

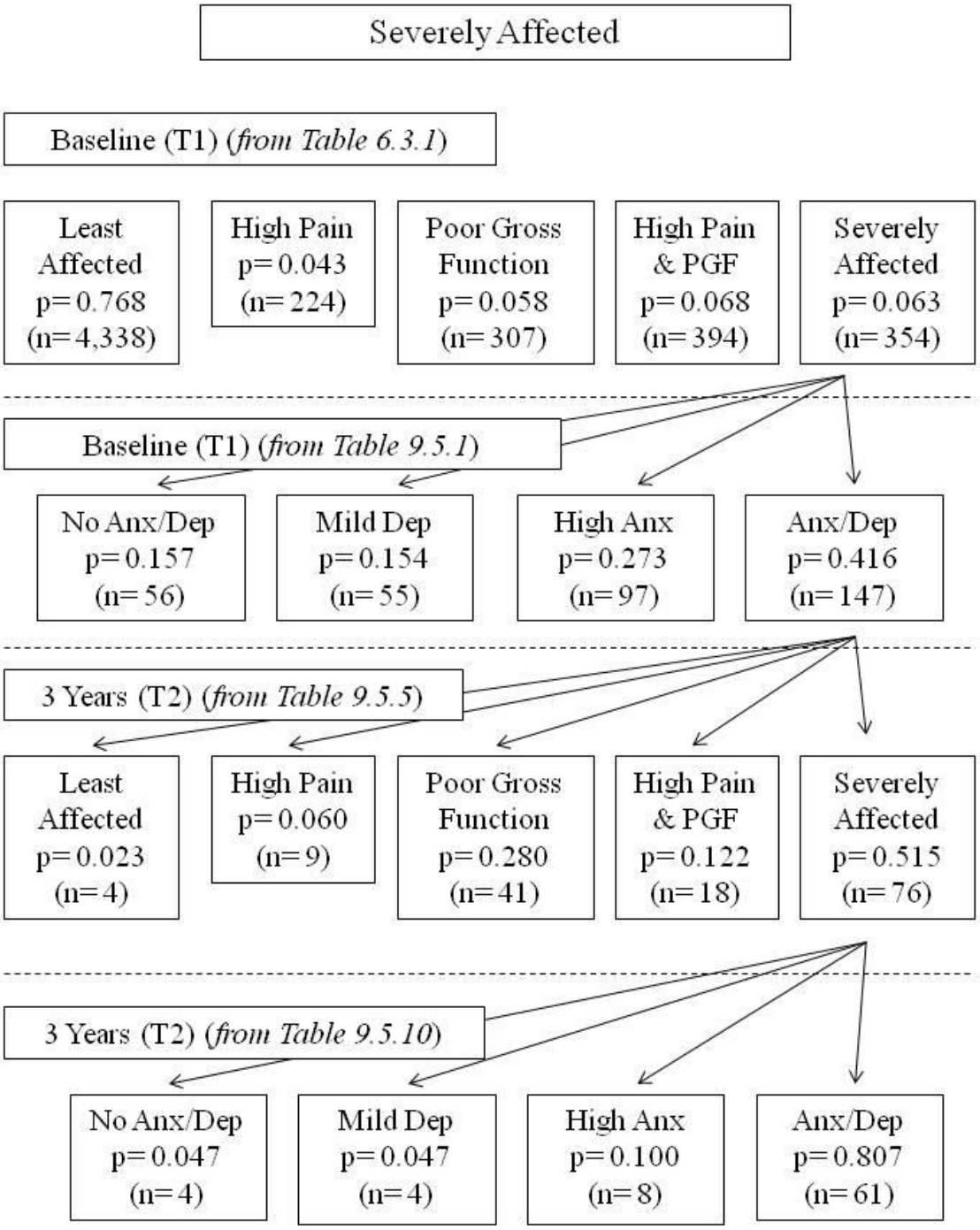
T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.0000	0.0000	1.0000	0.0000
High pain	0.0000	1.0000	0.0000	0.0000
Poor gross function	0.0000	0.0000	0.6138	0.3862
High pain & poor gross function	0.0000	1.0000	0.0000	0.0000
Severely affected	0.0000	0.0000	**0.7170**	0.2830

In Time 2 Severely affected, Time 2 Anx & dep:

T3 MH, on T3 hand, T2 MH, T2 hand	No anx/ dep	Mild dep	High anx	Anx & dep
Least affected	0.3098	0.0000	0.3733	0.3169
High pain	0.0000	1.0000	0.0000	0.0000
Poor gross function	0.0799	0.9201	0.0000	0.0000
High pain & poor gross function	0.0000	0.0875	0.2662	0.6463
Severely affected	0.0000	0.0000	0.0559	**0.9441**

Appendix K: Decision tree for severely affected

Appendix K.1: Flow diagram reflecting individuals estimated to be in the 'severely affected' hand state and 'anxiety and depression' at each time point (for baseline to 3 years).



Footnote: N.B. Observations (n) are estimated from probabilities (with the exception of hand states at baseline which are observed frequencies).

In *Figure L1*, the focus is on the relationship between the 'severely affected' hand state and the 'anxiety & depression' mental health state. The first line uses the observed frequencies from the item-response restricted LTA base model, with 354 individuals in the 'severely affected' state at baseline. Based on the concurrent relationship on individuals in 'severely affected' (*Table 9.5.1*), 0.416 were estimated to be in 'anxiety & depression', which equates to n=147 individuals (rounded to nearest integer). At the next stage, of individuals estimated to be in the previously mentioned states at baseline, 0.515 were estimated to remain in 'severely affected' at 3 years, which is n= 76 individuals (*Table 9.5.5*). Finally, 0.807 of these were estimated to remain in 'anxiety & depression' at 3 years (*Table 9.5.10*), estimating 61 individuals to follow this course over 3 years.