Gout flares and health-related quality of life in people living with gout: a prospective cohort study in primary care

Lorraine Watson

A thesis submitted for the degree of Doctor of Philosophy

October 2020

Keele University

Declaration

This PhD thesis is nested in a prospective cohort study at the Primary Care Centre Versus Arthritis, School of Primary, Community and Social Care, Keele University. Dr Edward Roddy developed the initial research idea and PhD proposal for this thesis. Cross-sectional analysis of health-related quality of life data from this cohort was presented in a PhD thesis by Dr Priyanka Chandratre supervised by Dr Edward Roddy. Dr Sara Muller undertook data cleaning and dealt with data requests to access the data included in this thesis.

The running of the cohort study involved staff at participating GP practices, West Midlands North Clinical Research Network, Keele University's Clinical Trials Unit and administrative & health informatics staff at Keele University's Primary Care Centre Versus Arthritis.

I confirm that I have planned and conducted all the analyses presented in this thesis. I have developed and written the research project under the guidance of my supervisors.

All supervisors (Dr Edward Roddy, Dr Elaine Nicholls, Dr John Belcher) provided feedback and guidance during the development and writing of this thesis. Statistical supervisors (Dr Elaine Nicholls, Dr John Belcher) provided guidance on the statistical analysis of data.

Acknowledgements

Firstly, I would like to thank my lead supervisor Dr Edward Roddy for his ongoing guidance, support and encouragement whilst I have been undertaking this thesis. I am extremely grateful for the support he has provided to both my academic and professional development.

I would also like to thank my current statistical supervisor Dr Elaine Nicholls and previous statistical supervisor Dr John Belcher for their guidance, support and encouragement. Their insight and reassurance whilst navigating the statistical analysis in this thesis has been invaluable.

Thank you to Prof Christian Mallen for his ongoing support and encouragement, including valuable feedback on presentations and manuscripts relating to this thesis. I would also like to thank Dr Sara Muller for her support during my thesis including the data cleaning, data requests and feedback on manuscripts.

I would like to thank the staff involved in the cohort study this thesis is nested in including staff at participating GP practices, West Midlands North Clinical Research Network, Keele University's Clinical Trials Unit and administrative & health informatics staff at Keele University's Primary Care Centre Versus Arthritis. Thank you to Dr Priyanka Chandratre for all her work establishing and running the cohort study.

I am very grateful to the Keele University Research User Group (RUG) for their involvement in this study and for sharing their views on both the thesis findings, implications and dissemination.

Importantly I would like to thank the participants in this study, without whom this thesis would not have been possible.

Finally, thank you to my family and friends for their fantastic support whilst undertaking this thesis.

Abstract

Background

Gout is a common inflammatory arthritis affecting an estimated 2.5% of the UK population. The hallmark of gout is sudden onset, extremely painful acute inflammatory flares, which are strongly associated with impaired health-related quality of life (HRQOL) in cross-sectional studies. Little is known about change in gout flares and HRQOL over time, and the factors associated with worse outcomes.

Methods

In a three-year prospective cohort study in primary care, people registered with gout reported gout flares and completed the Gout Impact Scale (GIS) subscales, 36-Item Short-Form Health Survey Physical Function subscale (SF-36 PF10) and Health Assessment Questionnaire Disability Index (HAQ-DI) at baseline, 6, 12, 24 and 36 months. Responders to the follow-up surveys and non-participants were compared to investigate attrition bias. Latent class growth analysis (LCGA) was used to identify and describe distinct gout flare trajectory classes. Linear mixed models were used to describe the factors associated with change in GIS subscales (disease-specific HRQOL), SF-36 PF10 and HAQ-DI (generic HRQOL).

Results

Responders had slightly less frequent gout flares and better HRQOL at baseline compared with non-participants. Six distinct gout flare trajectory classes were identified. The infrequent flare class had the lowest mean serum urate level and the highest proportion of participants taking allopurinol, whilst frequent flare classes had more participants who were socioeconomically deprived, obese, and had chronic kidney disease and oligo/polyarticular flares. Factors associated with deterioration in both disease-specific and generic HRQOL included more frequent gout flares, oligo/polyarticular flares, using allopurinol, body pain, worse pain severity, and worse depression score.

Conclusion

Gout flare trajectory classes with distinct characteristics and the factors associated with change in HRQOL were identified in people living with gout in primary care, highlighting people at risk of worse outcomes over three years and at greatest need of targeted interventions.

Abbreviations

ACR	American College of Rheumatology
AF	Atrial fibrillation
AIC	Akaike Information Criteria
AIMS	Arthritis Impact Measurement Scales
AMD	Age-related macular degeneration
ANZHT	Aotearoa New Zealand Health Tracker dataset
BC	Before Christ
BIC	Bayesian Information Criteria
BLRT	Bootstrap Likelihood Ratio Test
BLUPs	Best linear unbiased predictors
BMI	Body Mass Index
BSR	British Society of Rheumatology
CCI	Charlson comorbidity index
CDA	Concern during an attack
CFI	Comparative Fit Index
CHD	Coronary heart disease
CHF	Congestive heart failure
CI	Confidence interval
CKD	Chronic Kidney Disease
CO	Concern overall
CPRD	Clinical Practice Research Datalink
CRN	Clinical Research Network
CT	Computerised tomography
CVA	Cerebrovascular accident
CVD	Cerebrovascular disease
DASH	Dietary approaches to hypertension
DECT	Dual-enery computed tomography
eGFR	Estimated Glomerular Filtration Rate
EULAR	European League Against Rheumatism
EQ-5D	EuroQol 5 Dimensions
EQ-5D-3L	EuroQol 5 Dimensions 3 Levels
ES	Effect size
EU	European Union
GAD	Generalise Anxiety Disorder
GAD-7	Generalise Anxiety Disorder-7
GAQ	Gout Assessment Questionnaire
GIS	Gout Impact Scale
GMM	Growth Mixture Modelling
GP	General Practice
GWAS	Genome-wide association studies
HAQ-DI	Health Assessment Questionnaire Disability Index
HAQ-II	Health Assessment Questionnaire II
HL	Hyperlipidaemia
HPFS	Health Professionals Follow Up Study
HRQOL	Health-Related Quality of Life
HR	Hazard Ratio
HT	Hypertension
HV	Hallux Valgus
ICC	Intraclass correlation coefficient

IHD	Ischaemic heart disease
IPQ-R	Revised Illness Perception Questionnaire
IQR	Interquartile range
IMD	Indices of Multiple Deprivation
LCGA	Latent Class Growth Analysis
LGCM	Latent Growth Curve Model
LL	Log likelihood
LMM	Linear Mixed Models
LMR-LRT	Lo Mendell Rubin Likelihood Ratio Test
LRT	Likelihood Ratio Test
MCID	Minimal clinically important difference
MI	Myocardial infarction
MLE	Maximum likelihood estimation
MSE	Medication side effects
MOS	Medical Outcomes Study
MOS-20	Medical Outcomes Study Short Form Health Survey – 20
MRFIT	Multiple Risk Factor Intervention Trial
MSU	Monosodium urate
MTPJ	Metatarsophalangeal joint
NB	Negative Binomial
NHANES	National Health and Nutrition Examination Survey
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NRES	National Research Ethics Service
NR	Not reported
NRS	Numerical rating scale
NSAIDs	Non-steroidal anti-inflammatory drugs
OA	Osteoarthritis
OMERACT	Outcome Measures in Rheumatology Clinical Trials
OR	Odds ratio
OSA	Obstructive sleep apnoea
PCCVA	Primary Care Centre Versus Arthritis
PF10	Physical Function 10
PIS	Participant information sheet
PPIE	Patient public involvement and engagement
PHQ-9	Patient Health Questionnaire-9
PVD	Peripheral vascular disease
QOL	Quality of life
REC	Research and Ethics Committee
REML	Residual maximum likelihood estimation
RF	Renal failure
RMSEA	Root mean square error of approximation
RR	Relative risk
RUG	Research user group
SA	Sleep apnoea
SD	Standard deviation
SES	Socio-economic status
SF-36	36-Item Short-Form Health Survey
SF-20	20-Item Short Form Health Survey
SF-12	12-Item Short Form Health Survey
SRM	Standardised response mean

SU	Serum urate
SUA	Serum uric acid
THIN	The Health Improvement Network
TIA	Transient Ischaemic Attack
T2T	Treat-to-target
TLI	Tucker-Lewis index
UK	United Kingdom
ULT	Urate-lowering therapy
US	United States
UTN	Unmet treatment need
VAS	Visual analogue scale
WBDA	Wellbeing during an attack
WHO	World Health Organisation
WHO-QOL BREF	World Health Organisation Quality of Life Brief Version
WRMR	Weighted root-mean-square residual
ZIP	Zero-inflated Poisson

Glossary of terms

Gout flare	A clinically evident episode of acute inflammation induced by monosodium urate crystals (Bursill et al 2019), characterised by joint pain, tenderness, swelling, erythema and heat (Roddy, Mallen & Doherty, 2013; Taylor et al 2015; Teng, Nair & Saag, 2006).
Health-related quality of life (HRQOL)	A multi-dimensional concept, the dimensions of which can include physical functioning, social networks or support, psychological or emotional wellbeing, role functioning or performance, and general health (Bowling, 2005; Fitzpatrick et al 1992; Hennessy et al 1994; Testa & Simonson, 1996; Wilson & Cleary, 1995).
Disease-specific HRQOL measure	A HRQOL measure which assesses domains deemed more relevant to a disease of interest (Hickey et al 2005; Testa & Simonson, 1996). For example, the Gout Impact Scale (GIS) as a HRQOL measure of gout-specific HRQOL.
Generic HRQOL measure	A HRQOL measure which assesses domains applicable to a variety of health states (Hickey et al 2005; Testa & Simonson, 1996). For example, 36-Item Short-Form Health Survey (SF-36), Health Assessment Questionnaire Disability Index (HAQ-DI).
Minimal clinically important difference (MCID)	The smallest difference in a score on a domain of interest (e.g. quality of life) which patients perceive as beneficial (Jaeschke, Singer & Guyatt, 1989).
Trajectory	The within-person pattern of a longitudinal outcome over time (Curran, Obeidat & Losardo, 2010; Nagin & Odgers, 2010).
Urate-lowering therapy (ULT)	Medication used in the management of gout to lower serum urate levels e.g. Allopurinol, Febuxostat.
Read code	Alphanumeric medical codes which allow primary care health care workers to code clinical consultations in the UK (Chisolm, 1990; Porcheret et al 2004).

External communication of thesis

Oral presentations

May 2019	*Factors Associated with Change in Health-Related Quality of Life in People with Gout: A Three-Year Prospective Cohort Study in Primary Care* Royal College of General Practitioners Midland Faculty: Annual Education, Research and Innovation Symposium 2019, Staffordshire.
November 2018	*Gout Attack Trajectories in a 3-year Cohort Study in Primary Care* Midland Rheumatology Society, Wolverhampton. (oral presentation runner up prize award)
November 2018	*Gout Attack Trajectories In A 3-year Cohort Study In Primary Care* School for Primary Care Research Showcase, London.
September 2018	*Gout Attack Trajectories Identified In A 3-year Cohort Study In Primary Care* UK-RiME Showcase, Nottingham.
May 2018	*Gout attack trajectories in primary care: a three-year prospective cohort study* Royal College of General Practitioners Midland Faculty: Annual Education, Research and Innovation Symposium 2018, Staffordshire.

Poster presentations

May 2019	*Gout Attack Trajectories in a 3-year Cohort Study in Primary Care* British Society of Rheumatology Annual Conference, Birmingham.
March 2019	*Gout Flare Trajectories in a 3-year Cohort Study in Primary Care* Midlands Academy of Medical Sciences Research Festival, Birmingham.
September 2018	*Gout Attack Trajectories Identified In A 3-year Cohort Study In Primary Care* UK-RiME Showcase, Nottingham.
Published abstract	Watson, L., Belcher, J., Mallen, C.D. & Roddy, E. (2019) 202 Gout attack trajectories in a 3-year cohort study in primary care. *Rheumatology.* 58 (3), kez107.018. doi:10.1093/rheumatology/kez107.018

Table of Contents

List of figures

List of tables

Appendices

1 Chapter One Introduction and background

1.1 Overview of chapter and aim

The aim of this chapter is to introduce the topic of this thesis and to provide relevant background to the topic. This thesis relates to the analysis of outcomes reported by people living with gout in a prospective three-year cohort study in primary care. Thus, background relating to gout, gout flares, and health-related quality of life is presented.

1.2 Gout in history and society

Evidence of gout dates back to 7000 BC with pathological findings identified in Egyptian remains (MacKenzie, 2015). Early clinical descriptions of gout are attributed to Hippocrates (MacKenzie, 2015; Nuki & Simkin, 2006), and the writings of the Greek Physician Aretaeus referred to the classical presentation of gout in the first metatarsophalangeal joint (MTPJ) as 'podagra' ('foot trap') (Porter & Rousseau, 1998). The term gout is thought to derive from 'gutta', the Latin word for drop, referring to the historical belief that 'humours' of the body had dropped into a joint (Nuki & Simkin 2006; Porter & Rousseau, 1998). Thomas Sydenham described the clinical presentation of both acute and chronic gout, and how gout differs from other forms of arthritis, in as early as the 17th century (MacKenzie, 2015; Porter & Rousseau, 1998). Through history, gout has frequently been referred to as the 'disease of kings', due to historical associations with royalty and affluence (Nuki & Smikin, 2006) and those afflicted with gout have frequently endured caricature and a satirical portrayal (Porter & Rousseau, 1998). A recent content analysis of US & UK popular newspaper articles by Duyck, Petrie & Dalbeth (2016) revealed that gout still remains a source of humour and is frequently portrayed as a self-inflicted condition caused by overindulgence. Societal misconceptions regarding the cause and management of gout, which are also prevalent in healthcare professionals (Spencer, Carr & Doherty, 2012), can lead to stigma for people living with gout (Chandratre et al 2016; Lindsay et al 2011).

1.3 Pathophysiology, clinical manifestation and diagnosis of gout

Hyperuricaemia

Uric acid is a weak acid, which exists predominantly in its ionic form urate at physiological pH (Choi, Mount & Reginato, 2005; Mandal & Mount, 2015), and is the final breakdown product of purine metabolism in humans, produced through the action of xanthine oxidase on xanthine (Hediger et al 2005; Roch-Ramel & Guisan, 1999; Teng, Nair & Saag, 2006).

Underexcretion (inadequate removal) of urate from the body accounts for around 90% of hyperuricaemia (Choi, Mount & Reginato, 2005; Teng, Nair & Saag, 2006). Renal underexcretion is the predominant cause of hyperuricaemia (Dalbeth, Merriman & Stamp, 2016) as two-thirds of daily urate excretion is via the kidneys and one-third is excreted via the gut (Hediger et al 2005; Roch-Ramel & Guisan, 1999). After glomerular filtration of urate, renal transporters and exchanger molecules located in the proximal renal tubule play a key role in regulating serum urate levels; by the reuptake of urate (via SLC2A9, URAT1, OAT4 & OAT10) and urate secretion (via ABCG2, NPT1, OAT1-3, MRP4) (Dalbeth, Merriman & Stamp, 2016; Enomoto et al 2002; Hediger et al 2005; Mandal & Mount, 2015; Roch-Ramel & Guisan, 1999; Teng, Nair & Saag, 2006) see Figure 1.1. A range of factors, associated with reduced excretion of urate, predispose to hyperuricaemia and gout including chronic kidney disease, hypertension, loop or thiazide diuretics, obesity and genetic factors (Hyndman, Liu & Miner, 2016; Merriman, Choi & Dalbeth, 2014; Roddy, Mallen & Doherty, 2013) (see section 1.5).

In contrast to decreased excretion, an increase in endogenous (due to de novo synthesis or increased cell turnover) and exogenous sources of purines (due to dietary intake) accounts for a smaller proportion of elevated urate levels (Choi, Mount & Reginato, 2005; Dalbeth, Merriman & Stamp, 2016; Teng, Nair & Saag, 2006).

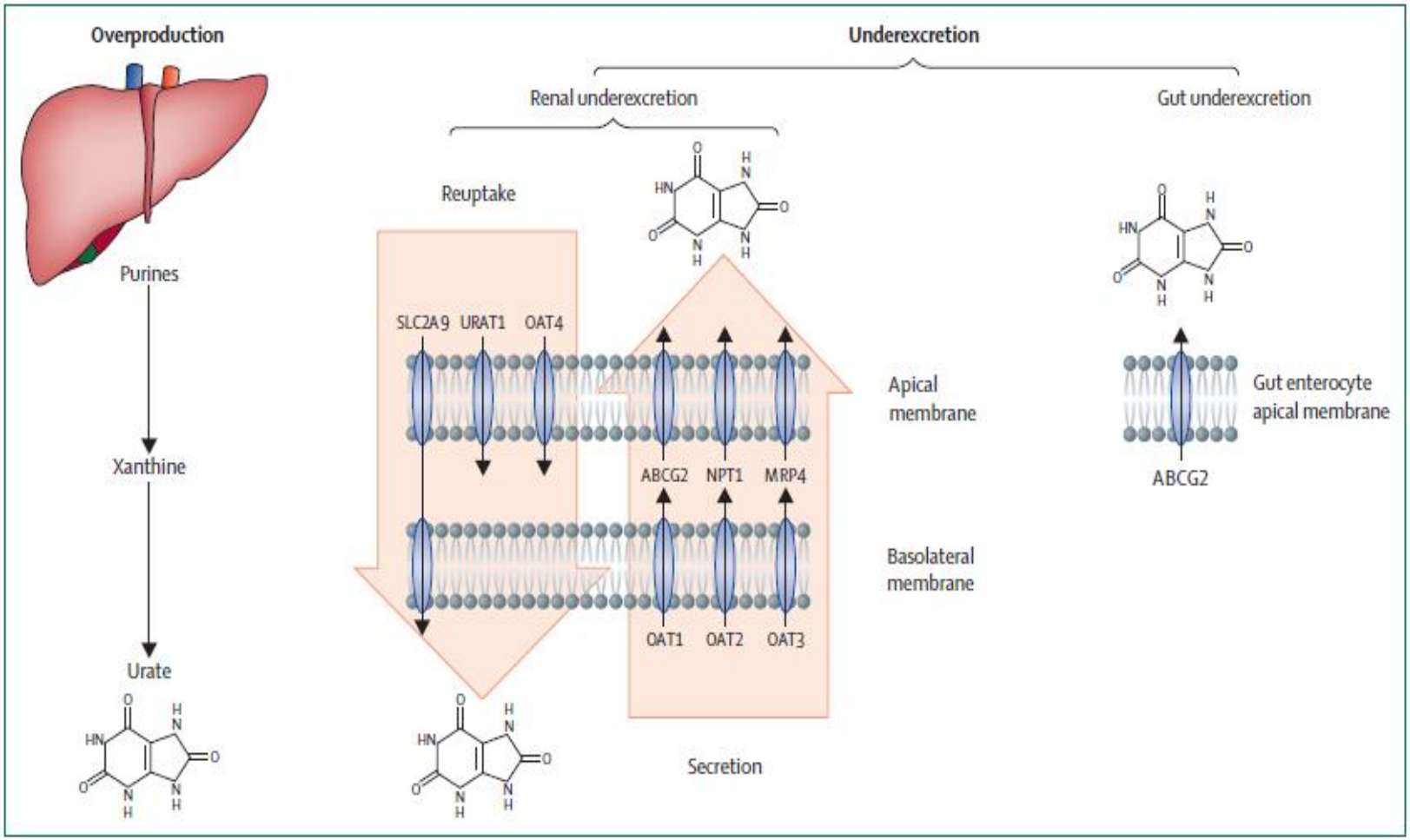


Figure 1.1 Mechanism of hyperuricaemia by Dalbeth, Merriman & Stamp (2016)

(included with permission from Elsevier Ltd)

Crystal formation

Elevation of serum urate (above 380 µmol/L [6·8 mg/dL]) leads to precipitation, nucleation and deposition of needle-shaped monosodium urate (MSU) crystals (Dalbeth, Merriman & Stamp, 2016; Roddy, Mallen & Doherty, 2013). Factors asserted to be associated with uric acid crystallisation and deposition include temperature, trauma, cartilage damage and pre-existing osteoarthritis (Choi, Mount & Reginato, 2005; Roddy, Zhang & Doherty, 2007a; Roddy, 2011).

Gout flares

When MSU crystals shed from their location in cartilage into the joint space, this can initiate the inflammatory process observed in a gout flare (Roddy, Mallen & Doherty, 2013). This inflammatory process involves the phagocytosis of MSU crystals by macrophages, which leads to the engagement of the inflammasome NALP3 (Martinon et al 2006; Petrilli & Martinon, 2007). The enzyme Capase-1 is activated by the engagement of NALP3 which leads to the production of biologically active pro-inflammatory cytokine interleukin-1β (IL-1β), expediting the arrival of neutrophils and other inflammatory cells (Dalbeth, Merriman & Stamp, 2016; Martinon et al 2006; Petrilli & Matrinon, 2007). The influx of inflammatory cells such as neutrophils, and the subsequent inflammatory response, results in synovitis and the clinical manifestations observed in a gout flare (Choi, Mount & Reginato, 2005; Martinon et al 2006; Petrilli & Matrinon, 2007). These symptomatic manifestations of flares involve excruciating joint pain and tenderness with associated swelling, erythema and heat (Roddy, Mallen & Doherty, 2013; Taylor et al 2015; Teng, Nair & Saag, 2006). The rapid onset of such symptoms is distinctive of a gout flare, with symptoms typically reaching peak intensity within the first 24 hours and resolving in around one to two weeks (Roddy, Mallen & Doherty, 2013; Taylor et al 2015). The first MTPJ is the most commonly affected joint (Roddy, 2011; Stewart et al

2016), followed by the midfoot, ankle, knee and fingers (Roddy, Zhang & Doherty, 2007a; Roddy, Mallen & Doherty, 2013; Taylor et al 2015). Monoarthritis is most commonly observed during a gout flare, however oligo- or polyarthritis (involvement of two to five joints or more than five joints respectively) may occur particularly as the disease process progresses (Dalbeth, Merriman & Stamp, 2016; Prowse et al 2013). Whilst the terms gout flare and gout attack have been used interchangeably, a recent consensus statement has endorsed the term gout flare to describe a clinically evident episode of acute inflammation induced by monosodium urate crystals (Bursill et al 2019).

Intercritical period

The natural history of gout includes periods of time without any overt symptoms of a flare; these periods are of a variable length of time and are referred to as the intercritical period (Dalbeth, Merriman & Stamp, 2016; Roddy, 2011; Roddy, Mallen & Doherty, 2013; Teng, Nair & Saag, 2006). Ultrasound studies have revealed that subclinical synovitis can persist in the intercritical period despite a lack of symptoms (Roddy et al 2013).

Tophi and chronic gout

The indicative characteristic of chronic gout, tophaecous deposits, are typically generated over many years via a chronic inflammatory pathway (Dalbeth et al 2010; Dalbeth, Merriman & Stamp, 2016) and are composed of an MSU crystalline core surrounded by zones of both pro-inflammatory and anti-inflammatory factors (Chhana & Dalbeth, 2015). Tophi present as nodular deposits of MSU crystals in subcutaneous tissue, which mainly occur in the toes, fingers, olecranon bursae, Achilles tendon, knees and the helix of the ear (Chhanna & Dalbeth, 2015; Roddy, Mallen & Doherty, 2013). Whilst tophi may initially be asymptomatic, advanced chronic tophaecous and arthritic gout can lead to chronic joint pain, structural damage, stiffness and limited activity (Dalbeth, Merriman & Stamp, 2016; Roddy, 2011).

Diagnosis

The gold standard for diagnosis of gout is the identification of MSU crystals in synovial fluid or tissue samples using polarising light microscopy (Dalbeth, Merriman & Stamp, 2016; Hui et al 2017; Richette et al 2020). Imaging, such as ultrasound, plain radiographs and dual-energy computed tomography (DECT), may also be used to aid the diagnosis of gout (Dalbeth, Merriman & Stamp, 2016; Hui et al 2017). However, synovial fluid microscopy or imaging are not routinely used to diagnose gout in primary care and a clinical diagnosis of gout can be made based on classic clinical manifestations of flares when synovial fluid analysis is not feasible (Richette et al 2020).

1.4 Epidemiology of gout

1.4.1 Prevalence

Gout is a common inflammatory arthritis with a reported prevalence in the UK of 2.5% based on a Read coded (codes for clinical consultations) diagnosis in the Clinical Practice Research Datalink (CPRD) (Kuo et al 2015a). Prevalence is higher in males (3.97%), compared with females (1.05%) (Kuo et al 2015a). Whilst Kuo et al (2015a) reported that gout was uncommon in individuals younger than 20 years (5.11 cases per 100 000 individuals), the prevalence increased with age for both males and females until the age of 80 years. A lower prevalence of gout of 1.4% was reported previously in CPRD records in the UK between 1990 and 1999 (Mikuls et al 2005) and in a GP database in the UK between 2000 and 2005 (Annemans et al 2008). Kuo et al (2015a) confirmed that the prevalence of gout in the UK had increased by 63.9% between 1997 and 2012.

Differences in study populations and methodologies between studies mean that the comparison of prevalence between countries is difficult (Kuo et al 2015b). Globally a higher prevalence of gout is observed in more developed countries and also in Pacific countries (Kuo

et al 2015b). A prevalence of gout of 3.8% in the US, based on a self-reported history of a gout diagnosis in participants in the National Health and Nutrition Examination Survey (NHANES) in 2007-2008, has been reported by Juraschek, Miller & Gelber (2013). Similar prevalences were identified more recently in NHANES 2015-2016 (Singh, Lingala & Mithal, 2019) and using medical record diagnoses in British Colombia, Canada (Rai et al 2017a). The prevalence of gout in Australia, based on general practice electronic records between 2008 and 2013, was reported to be 1.5% (Robinson, Taylor & Dalbeth, 2015). However, this is likely to be an underestimate as data for territories with more Australian Aborigines were not available (Robinson, Taylor & Dalbeth, 2015). Winnard et al (2012) reported a higher prevalence of gout in New Zealand of 2.7% in the Aotearoa New Zealand Health Tracker (ANZHT) dataset, with a higher prevalence of gout amongst Maori and Pacific populations. A higher prevalence of gout of 6.2%, based on health insurance database records, has been reported in Taiwan and gout was more prevalent in areas of Taiwan with greater aboriginal populations (Kuo et al 2015c).

1.4.2 Incidence

The incidence of gout reported by Kuo et al (2015a) in the UK in CPRD was 1.77 per 1000 person-years in all ages, with a higher incidence in men (2.58 per 1000 person-years) in comparison with females (0.99 per 1000 person-years). The incidence of gout had increased by 29.6% between 1997 and 2012 (Kuo et al 2015a). Incidences of gout in the UK were reported of 1.31 per 1000 person-years in CPRD between 1990 and 1999 (Mikuls et al 2005) and 2.68 per 1000 person-years in The Health Improvement Network (THIN) primary care database between 2000 and 2007 (Cea Soriano et al 2011). Whilst an incidence of 2.9 per 1000 person-years in Canada (Rai et al 2017a) and 2.74 per 1000 person-years in Taiwan have been reported (Kuo et al 2015c), a lower incidence of 0.95 per 1000 person-years was reported

in Italy (Trifiro et al 2013). The incidence of gout has been shown to increase with age; with a peak in incidence observed at 80-84 years (Kuo et al 2015a).

1.5 Factors associated with gout

1.5.1 Serum urate

Hyperuricaemia is a well-established risk factor for gout (Roddy & Choi, 2014). In population-based studies, higher levels of serum urate have been associated with a greater prevalence of gout (Campion, Glynn & Delabry, 1987; Lin, Lin & Chou, 2000). A recent study, using data from four cohorts, highlighted serum urate as a strong concentration-dependent predictor of incident gout with a 15-year cumulative incidence of 49% in people with serum urate ≥10 mg/dL (Dalbeth et al 2018). Hazard ratios (HRs) of 2.7 and 6.4 for incident gout in men with serum urate levels between 361 to 405 µmol/L and >405 µmol/L respectively, compared to those with a serum urate less than 361 µmol/L, were reported in data from a Swedish screening program (Kapetanovic et al 2018).

However, the presence of hyperuricaemia does not guarantee that an individual will acquire gout, and many individuals in population studies with elevated urate levels did not go on to develop gout (Brauer & Prior, 1978; Campion, Glynn & Delabry, 1987; Dalbeth et al 2018; Lin, Lin & Chou, 2000).

1.5.2 Chronic kidney disease & nephrolithiasis

The association between gout and chronic kidney disease (CKD) is bidirectional, as CKD is a risk factor for gout but gout is also a risk factor for CKD (Choi & Roddy, 2014; Roughley et al 2018). In a case-control study using CPRD the odds of incident gout were 5.96 times higher in people with renal disease compared with those without renal disease and also patients with gout were three times more likely to develop renal disease compared with those without gout (Kuo et al 2016). The increased odds of CKD in people with gout (OR 2.41) was also reported

in a systematic review and meta-analysis (Roughley et al 2015). The increased risk of CKD in people living with gout may be attributable to a number of potential mechanisms including hyperuricaemia, chronic inflammation, or medications such as non-steroidal anti-inflammatory drugs (NSAIDs) (Roughley et al 2018).

A systematic review and meta-analysis also reported that gout was associated with self-reported lifetime nephrolithiasis (OR 1.77) (Roughley et al 2015).

1.5.3 Metabolic syndrome

Metabolic syndrome is a common comorbidity in people living with gout with a prevalence of 62.8% amongst people with gout in the third NHANES (Choi et al 2007). Several of the clinical conditions, such as obesity, diabetes, hyperlipidaemia and hypertension, which are considered to be entities of metabolic syndrome are associated with gout (Roddy & Choi, 2014).

Obesity has been highlighted in a recent systematic review and meta-analysis as a major risk factor for incident gout, with individuals with a BMI $\geq$ 30 kg/m^2 being over two times more likely to develop gout (relative risk (RR) of 2.24) compared with non-obese individuals (Evans et al 2018). It is thought that obesity promotes insulin resistance which can reduce renal urate excretion (Ter Maaten et al 1997).

In the THIN dataset, the incidence of gout was lower in people with diabetes (RR 0.67) compared with individuals without diabetes (Rodriguez, Cea Soriabo & Choi, 2010). Diabetes was also not identified as a risk factor for incident gout in a more recent CPRD study (Kuo et al 2016). The lack of a positive association between diabetes and gout may be attributable to the uricosuric effect of glycosuria or an impaired inflammatory response in diabetes (Rodriguez, Cea Soriabo & Choi, 2010; Roddy & Choi, 2014).

A recent systematic review and meta-analysis identified that individuals with hypertension were 2.11 (RR) times more likely to develop gout than those without hypertension (Evans et al 2018).

1.5.4 Cardiovascular diseases

A higher risk of incident cardiovascular diseases such as cerebral vascular disease (CVD) (HR 1.29), myocardial infarction (MI) (HR 1.16) and congestive heart failure (CHF) (HR 1.81) has been observed in people with a diagnosis of gout, compared with people without gout (Kuo et al 2016). In a retrospective cohort study using CPRD Clarson et al (2014) reported that whilst both men and women with gout were at an increased risk of any vascular event, any coronary heart disease (CHD), and peripheral vascular disease (PVD) compared with those without gout, a greater risk was observed in women with HRs of 1.25 for any vascular event, 1.25 for any CHD, and 1.89 for PVD in women with gout compared with women without gout. The association between cardiovascular diseases and gout may be mediated by endothelial dysfunction caused by elevated uric acid or inflammation (Kuo et al 2016). In a retrospective cohort study using Medicare data from the US a diagnosis of gout was associated with a higher risk of incident atrial fibrillation (AF) (HR 1.71) (Singh & Cleveland, 2018a).

1.5.5 Other medical conditions

A higher odds of incident gout in people diagnosed with hypothyroidism compared with those without hypothyroidism was reported by Kuo et al (2016) (OR 1.50), and also a later case-control study by Bruderer et al (2017) (OR 1.12). The associations observed between hypothyroidism and gout could be attributable to the effect of thyroid hormones on glomerular filtration rate (GFR) affecting serum urate levels (Bruderer et al 2017).

Osteoarthritis (OA) was associated with an increased odds of incident gout (OR 1.27), and an increased risk of OA has been reported following a diagnosis of gout (HR 1.45) (Kuo et al 2016).

The deposition of monosodium urate (MSU) crystals in osteoarthritic cartilage is thought to play a part in the link between the two conditions (Roddy, 2011).

An increased risk of gout has been identified in patients with obstructive sleep apnoea (OSA) compared to those without OSA, in retrospective cohort studies using data from CPRD (Blagojevic-Bucknall et al 2019) and THIN database (Zhang et al 2015). It is thought that hypoxia-induced nucleotide turnover leads to hyperuricaemia (Zhang et al 2015).

An increased risk of erectile dysfunction in people with gout, in comparison with controls, was identified in a CPRD cohort (Sultan et al 2017). A recent observational cohort study using medicate claims data in the US reported a higher risk of age-related macular degeneration (AMD) in older adults with gout in comparison to those without gout (Singh & Cleveland, 2018b). These findings may be attributable to the effects of hyperuricaemia on vascular structure and oxidative stress (Singh & Cleveland, 2018b; Sultan et al 2017).

1.5.6 Medications

In a retrospective case–control analysis using the General Practice Research Database in the UK, an increase in the odds of incident gout was observed in patients currently using loop diuretics, thiazide diuretics, and thiazide-like diuretics (ORs 2.64, 1.70, and 2.30 respectively) (Bruderer et al 2014).

An increased risk of gout in people with hypertension currently using specific antihypertension medications was detected in the THIN database; with RRs of 1.48 for β blockers, 1.24 for angiotensin converting enzyme inhibitors, and 1.29 for non-losartan angiotensin II receptor blockers (Choi et al 2012).

1.5.7 Alcohol

In a Health Professionals Follow Up Study (HPFS) prospective cohort study, Choi et al (2004a)reported a RR of 2.53 for incident gout in males consuming more than 50g of alcohol

per day, compared with consuming no alcohol. A dose-response between the amount of alcohol consumed and the risk of incident gout was observed (Choi et al 2004a). The association between alcohol and increased risk of gout may be attributable to increased purine intake and nucleotide turnover (beer is particularly high in the purine, guanosine), along with reduced renal excretion due to lactic acidosis (Roddy & Choi, 2014).

1.5.8 Diet

In the HPFS, a higher risk of gout also was reported in participants with higher intakes of meat and seafood (Choi et al 2004b), sugar sweetened soft drinks (Choi et al 2008), free fructose (Choi et al 2008) and also participants with higher 'western diet' pattern scores (characterised by red or processed meats, sugar sweetened drinks, sweets, desserts, and refined grains) (Rai et 2017b). Conversely, a higher consumption of dairy foods (Choi et al 2004b), vitamin C (Choi, Gao & Curhan, 2009), coffee (Choi, Willet & Curhan, 2007), decaffeinated coffee (Choi, Willet & Curhan, 2007) and higher 'DASH' (Dietary Approaches to Hypertension) dietary pattern scores (Rai et 2017b) were associated with a reduced risk of gout.

1.5.9 Genetics, serum urate and gout

Twin studies have revealed that an estimated 45-75% of serum urate levels can be explained by inherited genetic variants (Dalbeth, Stamp & Merriman, 2017; Major et al 2018). Observational genome-wide association studies (GWAS) have identified a range of genes coding for individual urate transporters and proteins associated with metabolic pathways (Dalbeth, Stamp & Merriman, 2017; Major et al 2018), with genes coding for the transporter SLC2A9 thought to explain 2 to 3 % of the variation in urate in European populations (Dalbeth, Stamp & Merriman, 2017; Major et al 2018).

In contrast to our knowledge of the genetic influence on urate, less is known about the role of genetics in the development of gout (Major et al 2018). Genes associated with gout which

encode proteins involved in the pathogenesis of gout flares, TLR4 gene and NLRP3 inflammasome gene, have been identified in candidate gene studies (Dalbeth, Stamp & Merriman, 2017; Major et al 2018). GWAS have indicated that variants of ABCG2 and SLC2A genes were associated with gout in European participants (Major et al 2018).

1.6 Management of gout flares

Colchicine or non-steroidal anti-inflammatory drugs (NSAIDs) are advocated as first line management of gout flares in the management guidelines from the British Society of Rheumatology (BSR) (Hui et al 2017) (see Figure 1.2), European League Against Rheumatism (EULAR) (Richette et al 2017), and American College of Rheumatology (ACR) (Khanna et al 2012a). Which medication is selected to manage acute gout flares will be influenced by patient choice, renal function, and existing comorbidities (Hui et al 2017; Khanna et al 2012a; Richette et al 2017).

1.6.1 Colchicine

Colchicine is a naturally occurring alkaloid which modulates both pro- and anti-inflammatory pathways involved in the pathogenesis of gout (Dalbeth, Lauterio & Wolfe, 2014; Roddy, Mallen & Doherty, 2013). Systematic reviews (Khanna et al 2014; Van Echteld et al 2014) report an improvement in pain following the use of colchicine during gout flares, but colchicine use was associated with gastrointestinal symptoms and other adverse events.

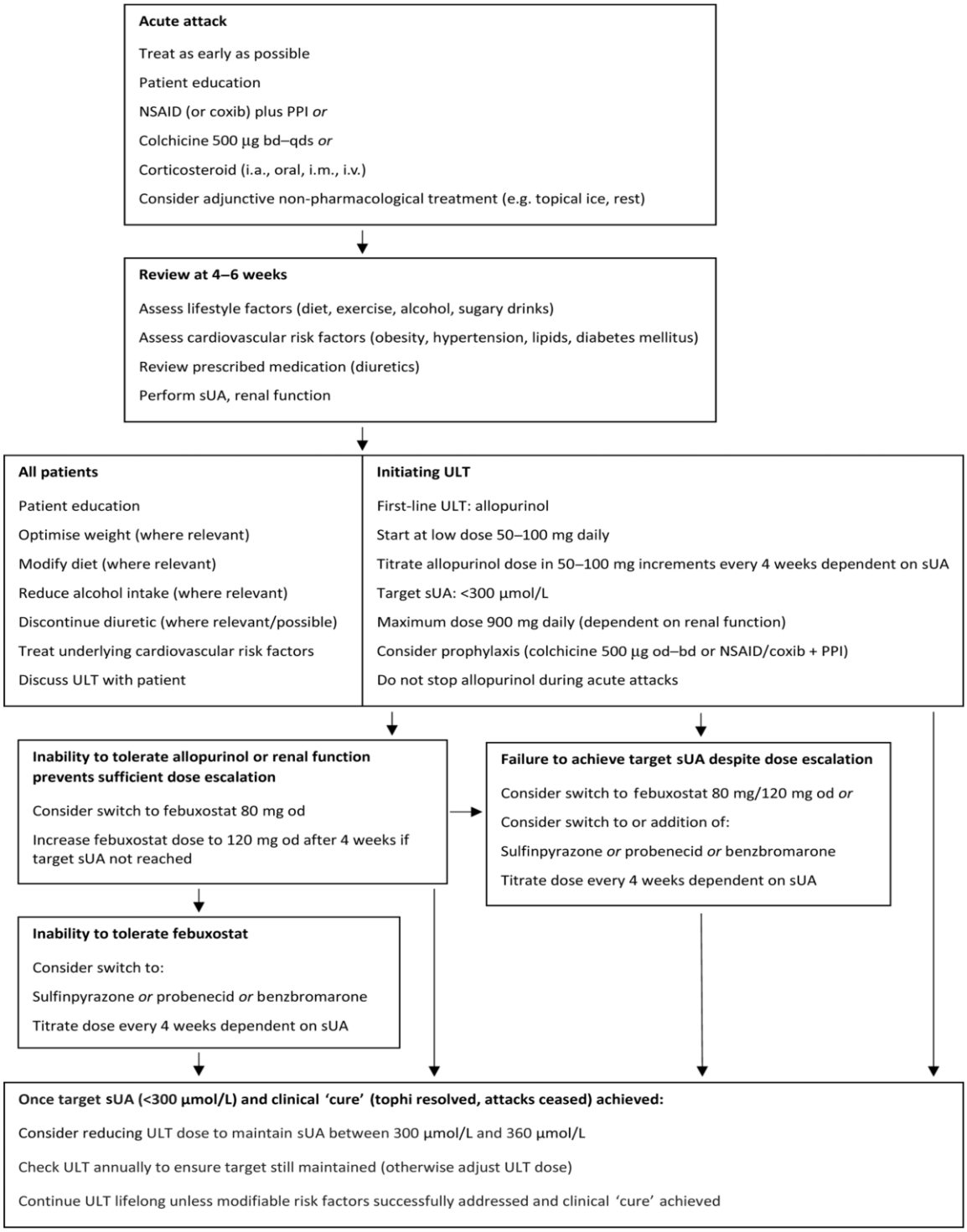


Figure 1.2 Algorithm for the management of gout by Hui et al (2017)

(included with permission from Oxford University Press)

Coxib, cyclooxygenase-2 inhibitor; PPI, proton pump inhibitor; sUA, serum uric acid

1.6.2 Non-steroidal anti-inflammatory drugs

Non-steroidal anti-inflammatory (NSAID) medication at full dose, with the co-prescription of a gastro-protective agent, has been advocated as first line acute management of gout flares in the BSR (see Figure 1.2), EULAR and ACR gout management guidelines (Hui et al 2017; Khanna et al 2012a; Richette et al 2017) as an alternative to colchicine. NSAIDs are more commonly prescribed than colchicine and prednisolone for gout flares in the UK (Mikuls et al 2005; Roddy et al 2010).

A systematic review of treatment strategies for acute gout by Khanna et al (2014) revealed 23 RCTs reporting the efficacy of NSAIDs (e.g. indomethacin, naproxen) in the management of flares. No difference in pain intensity was observed by Roddy et al (2019) in people randomised to either colchicine or naproxen treatment, but naproxen was associated with fewer side-effects.

NSAIDs are contraindicated in patients with impaired renal function or a history of gastrointestinal ulceration or perforation (Hui et al 2017).

1.6.3 Corticosteroids

Corticosteroids, oral or intra-articular injection, have been advocated as alternative treatment for gout flares by BSR (see Figure 1.2) and EULAR gout management guidelines for individuals who are unable to tolerate colchicine or NSAIDS (Hui et al 2017; Richette et al 2017). Oral prednisolone has been shown to be an equivalent treatment to NSAIDs (Janssens et al 2008; Man et al 2007). Whilst a systematic review by Wechalekar et al (2013) failed to identify any RCTs investigating the efficacy and safety of intra-articular glucocorticoids for gout flares, their use has been advocated for acute management, particularly monoarticular flares (Hui et al 2017).

1.6.4 Non- pharmacological management

Guidance to rest, elevate and ice a joint affected by a gout flare is based on expert opinion (Hui et al 2017). A Cochrane review of lifestyle interventions for acute gout by Moi et al (2013) found low quality evidence from a single study that the addition of topical ice therapy to oral prednisolone and colchicine for people experiencing an acute flare led to a greater pain reduction at one week.

1.7 Urate-lowering therapy and gout management

1.7.1 Allopurinol

Allopurinol is recommended as first line therapy in the management of gout (Hui et al 2017; Khanna et al 2012b; Richette et al 2017; Seth et al 2014) (Figure 1.2). Allopurinol is a purine non-specific xanthine oxidase inhibitor which inhibits the synthesis of uric acid (Abhishek, Roddy & Doherty, 2017; Roddy, Mallen & Doherty, 2013) and is later metabolised to oxypurinol which is renally excreted (Abhishek, Roddy & Doherty, 2017). A Cochrane review of 11 trials (RCTs and quasi RCTs) by Seth et al (2014) investigated the efficacy and safety of allopurinol in comparison to either placebos or active comparators (alternative medication or alternative different dose of allopurinol) concluding that, based on low to moderate quality evidence, allopurinol was more successful than placebo in achieving target serum urate levels and there were similar withdrawals compared to placebo and active comparators. A starting dose of 50-100 mg of Allopurinol has been advocated by practice guidelines, followed by up-titration of Allopurinol to achieve target serum urate levels (Hui et al 2017; Khanna et al 2012b; Richette et al 2017). A serum urate level below 360 µmol/L was achieved by 95% of participants receiving up titration of their allopurinol dose in a recent nurse-led treat to target (T2T) RCT by Doherty et al (2018).

The BSR gout management guidelines (Hui et al 2017) advocated that the option of ULT should be discussed with all patients with a diagnosis of gout, but ULT should be particularly advised if patients experience ≥2 gout flares in the previous 12 months, tophi, chronic gouty arthritis, joint damage, renal impairment, renal stones, a prescription for diuretics or a diagnosis of gout at a young age.

Allopurinol is usually well tolerated although around 2% of users may develop a rash and 5% may stop allopurinol due to an adverse event (Stamp et al 2012). BSR gout management guidelines recommend adjustment of the starting dose of Allopurinol depending on a patient's eGFR (Hui et al 2017).

Despite the availability of allopurinol gout is frequently sub-optimally managed and only a third of people with gout consulting a GP in the UK in 2012 received urate-lowering therapy (ULT) (Kuo et al 2015a). When ULT is prescribed, in the majority of patients the dose is not adjusted to optimise urate levels (Cottrell et al 2013), and over two thirds of patients prescribed ULT are non-adherent to ULT (Scheepers et al 2018).

1.7.2 Febuxostat

Febuxostat has been recommended by the National Institute for Health and Care Excellence (NICE) for ULT in situations where allopurinol is not tolerated or contraindicated, including where renal impairment precludes an increase in allopurinol dose to achieve target serum urate levels (Hui et al 2017; National Institute for Health and Care Excellence, 2008) (Figure 1.2). Febuxostat is a non-purine xanthine oxidase inhibitor which unlike allopurinol inhibits both the oxidised and reduced form of xanthine oxidase and is metabolised in the liver (Liu et al 2018; Roddy, Mallen & Doherty, 2018; Ye et al 2013).

1.7.3 Uricosuric agents

Uricosuric medication reduces serum urate levels by increasing the amount of uric acid which is secreted by the kidneys (Kydd et al 2014). Uricosuric medications such as sulfinpyrazone, probenecid and benzbromarone may be used as ULT where individuals are intolerant, resistant or have contraindications to xanthine oxidase inhibitors (Hui et al 2017; Richette et al 2017). Losartan and fenofibrate have mild uricosuric effect but should not be used as the sole ULT. They can be considered as adjunctive ULT where medication for hypertension or lipid lowering respectively is required (Hui et al 2017).

Allopurinol is the most commonly used ULT in UK general practice, with 38% of patients with gout prescribed allopurinol and less than 1% of patients prescribed febuxostat (0.2%) or uricosuric agents (0.1%) as a first line ULT between 1987 to 2014 (Scheepners et al 2018).

1.7.4 Diuretics and lifestyle factors

A review of diuretics prescribed to individuals with a diagnosis of gout is recommended and where possible alternatives to thiazide and loop diuretics can be prescribed (Hui et al 2017; Richette et al 2017).

Guidelines for the management of gout recommend that, where relevant, dietary modification, reduction in alcohol intake, weight optimisation/loss, and modification of cardiovascular risk factors should be promoted (Hui et al 2017; Khanna et al 2012b).

1.8 Gout flares

Rapid onset acute inflammatory gout flares, characterised by excruciating joint pain, tenderness, swelling, erythema and heat, are the clinical hallmark feature of gout (Roddy, Mallen & Doherty, 2013; Taylor et al 2015; Teng, Nair & Saag, 2006). More frequent flares are associated with poorer HRQOL (Hirsch et al 2008; Khanna et al 2012c; Scire et al 2013), functional disability (Lee et al 2009; Scire et al 2013), higher use of healthcare resources

(Jackson et al 2015; Khanna et al 2012c; Saseen et al 2012), and lower work productivity (Khanna et al 2012c).

Gout flares have been identified as a key patient-reported outcome measure in the management of chronic gout by both people living with gout and clinicians (Schumacher et al 2009). However, until recently there has been no widely agreed definition of gout flare used in the research literature and no consensus on how gout flares should be measured as an outcome (Gaffo et al 2012; de Latour, Dalbeth & Taylor, 2015; Taylor et al 2009). Consequently, a standardised definition of a gout flare outcome measure has yet to be endorsed as a clinical outcome measure by Outcome Measures in Rheumatology Clinical Trials (OMERACT) (de Latour, Dalbeth & Taylor, 2015). A recent definition of gout flares has been validated for use in clinical research (Gaffo et al 2018). This definition of gout flares requires fulfilment of three or more of the following criteria: a patient-defined gout flare, pain score of three or more out of ten, the presence of at least one swollen joint, or the presence of at least one warm joint (Gaffo et al 2018). However, the authors have acknowledged the potential limitations regarding the generalisability of the validation process for this definition to primary care (Gaffo et al 2018).

Where the frequency of flares has been investigated previously, flares have been recorded by various different methods and over different reference time-periods. For example, studies have reported either the mean, median or number (proportion) of gout flares in three months (Dalbeth et al 2011; Dalbeth et al 2013; Wallace et al 2016; Wu et al 2009), six months (Wu et al 2009), 12 months (Becker et al 2009; Khanna et al 2012c; ten Klooster et al 2014; Lee et al 2009; Rashid et al 2015; Saseen et al 2012; Scire et al 2013; Spaetgens et al 2015; Strand et al 2012; Wu et al 2009), or another length of time (Annemans et al 2008; Halpern et al 2009; Rothenbacher et al 2011; Sarawate et al 2006). Whilst some studies use self-reported gout flares (Dalbeth et al 2013; Khanna et al 2012c; Khanna et al 2015; ten Klooster et al 2014),

many studies use gout consultations or prescriptions recorded in health insurance, clinical consultation or prescription databases, assuming these were for gout flares (Annemans et al 2008; Halpern et al 2009; Jackson et al 2015; Rashid et al 2015; Rothenbacher et al 2011; Sarawate et al 2006; Saseen et al 2012; Wu et al 2009).

Observational studies which report the number of gout flares over a 12-month period, have reported a mean of 3.1-3.9 (ten Klooster et al 2014; Spaetgens et al 2015) and median of 1 (Scire et al 2013). However, these included patients recruited from secondary care Rheumatology clinics only. Of the participants in the US and EU Health and Wellbeing survey, a quarter reported experiencing no flares in the previous 12 months, whilst over a third reported three or more flares during the same period (Khanna et al 2012c). Conversely in a population-based survey by Proudman et al (2019) using data from the South Australian Health Omnibus Survey, most participants (58%) reported no flares in the last 12 months, but 25% of participants reported two or more flares in this time period.

1.8.1 Trajectories of clinical outcomes

Whilst existing studies have reported the number of flares experienced by people living with gout over different time-periods, little is actually known about the pattern or 'trajectory' of flares over time. Over the last decade, there has been increased interest in utilising latent class growth analysis (LCGA) to investigate trajectories in longitudinal cohorts (Jung & Wickrama, 2008; Kongsted et al 2016). LCGA has been used to identify trajectories of pain, activity limitation, psychological factors, and comorbidities in a range of musculoskeletal and rheumatological conditions (Holla et al 2014; Nicholls et al 2014; Rzewuska et al 2015; Strauss et al 2014; Verkleji et al 2012). Trajectories of other clinical parameters such as disease activity and pain have also been identified using other methods such as growth mixture modelling (GMM) (Ogollah et al 2018; Siemons et al 2014). However, methods such as LCGA or GMM have yet to be utilised to investigate latent trajectories of gout flares.

1.8.2 Factors associated with gout flares

Gout characteristics

Several studies have investigated the relationship between serum urate levels and risk of gout flares. In a retrospective review of medical claims data, Halpern et al (2009) showed that the odds of gout flare in a two-year period were 1.3 times higher in patients with a serum urate level over 6 mg/dL in comparison with patients with a serum urate below 6 mg/dL (OR 1.3). Further retrospective reviews of medical claims data reported that a serum urate level over 6 mg/dL or initiation of ULT was associated with a greater frequency of gout flares in 12 months (Jackson et al 2015; Sarawate et al 2006) and higher average number of gout flares was seen at higher serum urate levels (Wu et al 2009).

In a case-control study by Abhishek et al (2016), a serum urate greater than 541 μmoles/litre, and disease duration of 13.4 years or more was associated with an increased odds of more than 2 gout flares in a 12 month period (OR of 1.96 and 1.97 compared with a serum urate ≤ 480 μmoles/litre and disease duration ≤ 5.1 years respectively). Whilst in a retrospective cohort study of patients newly initiated on ULT, Rashid et al (2015) identified that the odds of gout flares in a 12 month period were higher in participants with a serum urate level greater than 6.0 mg/dl (OR 1.64 one to two flares, OR 3.61 three or more flares), compared with those with a serum urate ≤ 6.0 mg/dl.

Medications

The use of certain medications has been associated with an increased risk of gout flares. In an internet-based case-crossover study there was an increased odds of a recurrent flare (a flare within a 12-month period in people with existing gout) in participants who self-reported using aspirin (OR 1.81) (Zhang et al 2014) or diuretics (OR 3.6) (Hunter et al 2006) in the previous 48 hours, compared with those who did not. Participants in a retrospective cohort

with a prescription for a diuretic or anti-inflammatory medication in their medical record were more likely to have had gout flares in a 12-month period (OR of 1.23 for ≥3 flares with diuretic prescription, OR of 3.10 for ≥3 flares with anti-inflammatory medication prescription), compared with those without prescriptions for diuretic or anti-inflammatory medication (Rashid et al 2015). However, no difference was identified in the number of flares experienced in the previous 12 months in participants taking diuretics, compared with those not taking diuretics, in a case-control study by Mitnala et al (2016).

Comorbidities

Renal disease, hypertension and ischaemic heart disease were identified as risk factors for first post-diagnosis gout flare in the UK THIN database (Rothenbacher et al 2011) (HR of 1.33, 1.15, 1.12 respectively). The number of comorbidities was a risk factor for gout flares in a retrospective cohort study of patients newly initiated on ULT, and participants who had three or more comorbidities had an increased odds of 1-2 flares or ≥3 flares during a 12-month period (OR 1.34, 1.93 respectively) (Rashid et al 2015).

BMI, alcohol consumption and dietary factors

Obesity (BMI ≥30 kg/m^2) was identified as an independent risk factor for first post-diagnosis gout flare in the UK THIN database (Rothenbacher et al 2011) (HR 1.22). Proudman et al (2019) reported that a higher BMI was associated with a greater probability of flares in the South Australian Health Omnibus Survey, and in participants who reported flares there was a greater probability of two or more flares. Baseline BMI was not associated with recurrent gout flares in an analysis of observational data from the Multiple Risk Factor Intervention Trial (MRFIT) database, however more than a 5% increase or decrease in BMI was associated with an increase and decrease respectively in recurrent gout flares (OR 1.60, 0.61 respectively) compared with no change in BMI (Nguyen et al 2017). A systematic review of longitudinal

studies reporting weight loss in people with gout reported that six out of eight studies reported an improvement in gout flares (Nielsen et al 2017).

An internet-based case-crossover study by Zhang et al (2006) identified that, compared to no alcohol consumption, consuming alcohol was associated with an increased odds of recurrent gout flares (OR 2.5 for ≥7 drinks in the last 48 hours) and the relationship was present for all types of alcohol reported. A dose-response relationship between the amount of alcohol consumed in the previous 24 hours and the odds of a recurrent flare was also observed (Zhang et al 2006). Later internet-based case-crossover studies found that the highest quintile of purine intake (Zhang et al 2012a) and cherry consumption (Zhang et al 2012b) over a two-day period was associated with an increased and reduced odds of recurrent gout flares respectively (OR 4.76 compared with lowest quintile of purine intake, OR 0.65 compared with no cherry consumption). A randomised double-blind controlled trial by Dalbeth et al (2012) demonstrated a reduction in gout flares over a three-month period in participants receiving enriched skimmed milk powder.

Socio-demographic

Non-attendance at further education and living in a socioeconomically deprived area were associated with having two or more gout flares over a 12-month period in a cross-sectional analysis in primary care (Bowden-Davies et al 2018). Younger age but not socio-economic status (SES) was associated with flare frequency in the preceding year in a population-based study using data from the South Australian Health Omnibus Survey (Proudman et al 2019). This is in contrast to the study by Rashid et al (2015) which reported that patients over the age of 65 years were more likely to experience three or more flares (OR 1.81) compared with individuals aged ≤65 years.

In summary, serum urate, disease duration, medications, number or type of comorbidities, obesity, alcohol consumption, diet, educational attainment, socio-economic status and age

have been associated with recurrent flares or flare frequency. Although some previous studies have reported the risk of gout flares in a follow-up time period, there has been no studies using data from multiple follow-up time-points to identify and describe people experiencing different patterns or trajectories of gout flares over time.

1.9 Health-related quality of life (HRQOL)

Interest in health research relating to quality of life has increased significantly over the last 30 to 40 years (Hickey et al 2005; Fayers & Machin, 2016). Despite interest in health-related quality of life (HRQOL), there remains a lack of consensus regarding a precise definition of HRQOL and the terms 'quality of life', 'health-related quality of life', 'wellbeing' 'health status', 'functional status' and 'satisfaction' are often used interchangeably (Bakas et al 2012; Dijkers, 2007; Moons, 2004; Taillefer et al 2003). HRQOL is multidimensional and conceptualisations of HRQOL often include the dimensions of physical functioning, social networks or support, psychological or emotional wellbeing, role functioning or performance, and general health (Bowling, 2005; Fitzpatrick et al 1992; Hennessy et al 1994; Testa & Simonson, 1996; Wilson & Cleary, 1995). The various published definitions of HRQOL, displayed in Table 1.1, concur with the view that HRQOL is a multifaceted concept. The assessment of HRQOL can involve the use of generic HRQOL instruments; to assess a variety of domains applicable to a variety of health states, and disease specific instruments; to assess domains deemed more relevant to a disease of interest (Hickey et al 2005; Testa & Simonson, 1996). An individual's subjective assessment of their HRQOL may be influenced by their perception or satisfaction of their situation, in contrast to their expectations or aspirations (Bowling, 2003; Felce & Perry, 1995; Moons, Budts & De Geest, 2006).

Table 1.1 Published definitions of health-related quality of life

Authors	**Definition of health-related quality of life**
Bowling, 2005 p.7	"Quality of life, then, is about **the goodness** of life, and in relation to health it is about the **goodness of those aspects of life affected by health**".
Carr, Gibson & Robinson, 2003 p.10	"Health-related quality of life is a broader concept concerned with whether disease or impairment **limits one's ability to fulfil a normal role**".
Cieza & Stucki, 2008 p.311	"Health-Related Quality of Life can be defined as an **individual's perceptions** of health and health related domains of **well-being**".
Hays & Reeve, 2008 p.241	"Health-related quality of life (HRQOL) refers to how well a person **functions in their life** and his or her **perceived wellbeing** in **physical, mental, and social domains** of health".
Hennessy et al 1994 p.665	"Health-related quality of life is multidimensional and is composed of, at minimum, **physica**l **functioning, psychological well-being**, **social** and **role** functioning, and **health perceptions**".
Kempen, Jelicic & Ormel, 1997 p. 539	"Health-related quality of life (HRQL) is a broad, **multidimensional** concept covering significant domains of **daily functioning** and **subjective** experience, such as **physical functioning**, **role** and **social functioning**, somatic sensation, **perceived health**, and **subjective well-being**."
Moons, Budts & de Geest, 2006 p.896	"Health-related quality of life has been developed to describe aspects of an individual's **subjective experience** that relate both directly and indirectly to **health, disease, disability and impairment**…..and the effectiveness of treatment".
Muldoon et al 1998 p.542	"..study of the patient's personal morbidity- that is the various **effects** that illness and treatments have on **daily life** and **life satisfaction**".
Skevington, Lofty & OConnell, 2004 p.299	"An individual's perception of their **position in life** in the context of the **culture** and **value** systems in which they live, and in relation to their **goals, expectations, standards and concerns**".
Testa & Simonson, 1996 p. 835	"health-related quality of life refer to the **physical**, **psychological**, and **social** domains of health, seen as distinct areas that are influenced by a person's **experiences, beliefs, expectations, and perceptions**".

1.9.1 The importance of HRQOL measurement

Patient-completed HRQOL measures can provide information on the wider impact of a disease on the individual involved (Haywood et al 2015; Calvert & Freemantle, 2003; Moons, 2004). This information can include details relating to the prevalence of any dysfunction or disability, in addition to psychosocial or emotional problems (Calvert & Freemantle, 2003; Detmar et al 2002; Higginson & Carr, 2001). Self-completion of HRQOL measures by patients is important as healthcare professionals and patients judge quality of life from very different perspectives (Fayers & Machin, 2016; Fitzpatrick et al 1992). The use of HRQOL measures over a period of time can also help to detect the dynamic impact of a disease, due to psychosocial adaptation or changes in the course of the disease (Fayers & Machin, 2016; Hahn et al 2007).

Data obtained through HRQOL measurement have the potential to facilitate decision-making concerning clinical management, service development and health policy (Green, Brazier & Deverill, 2000; Hahn et al 2007; Hennessy et al 1994; Fayers & Machin, 2016), through an understanding of the wider impact of disease on patients and identifying patients at risk of poorer outcomes (Calvert & Freemantle, 2003; Fitzpatrick et al 1992).

1.9.2 HRQOL in people living with gout

The various instruments which have been used to assess HRQOL in people with gout were described, and their clinimetrics evaluated in a systematic review by Chandratre et al (2013). The instruments most commonly used to measure generic HRQOL were found to be the 36-Item Short-Form Health Survey (SF-36) (Ware & Sherbourne, 1992) and Health Assessment Questionnaire Disability Index (HAQ-DI) (Bruce & Fries, 2003), which have been endorsed by the OMERACT group as patient reported outcome measures in chronic gout (Singh et al 2011a). The newer disease-specific measure, the Gout Impact Scale (GIS) (Hirsch et al 2008), was used by a smaller number of studies (Chandratre et al 2013). More information relating to the interpretation of these instruments along with the validity, reliability and

responsiveness of these instruments can be found in chapter three. Other instruments which have been used to measure HRQOL in people with gout included AIMS, EQ-5D, HAQ-II, MOS 20, SF-12, WHO-QOL BREF (Chandratre et al 2013).

Studies which have investigated HRQOL in people living with gout which were reported in the systematic review by Chandratre et al (2013) and identified in a subsequent review of the literature for this thesis are listed in appendix 1.

Lower scores on the physical component of the SF-36, indicating worse generic HRQOL, have been reported in people living with gout in comparison with population normative values (Becker et al 2009; Khanna et al 2011a; Lee et al 2009; Strand et al 2012), and individuals matched on age and gender (Fu et al 2017; Scire et al 2013). Worse quality of life overall and on the physical domain, measured using the WHO-QOL BREF, was reported by Roddy, Zhang & Doherty (2007c) in people with gout registered with two general practices in the UK, in comparison with controls. More recently, worse physical component scores (PCS) on the SF-12 were reported in people who reported a diagnosis of gout in the population-based survey by Proudman et al (2019), in contrast with those without gout.

Conversely no difference between quality of life in the mental or psychological domains in people living with gout, in comparison with either controls or population normative values, has been reported by several studies (Becker et al 2009; Roddy, Zhang & Doherty, 2007c; Scire et al 2013; Singh & Strand, 2008). In fact some studies suggest that quality of life on the mental domain may in fact be slightly better for people living with gout; Lee et al (2009) reported that people living with gout over the age of 75 had a higher SF-36 mental component score (MCS) than the general population, whilst Khanna et al (2011a) reported that the SF-36 MCS for participants with gout were 0.2 SD above the US population norm.

The level of activity limitation and disability, measured via HAQ-DI, in people living with gout reported in observational studies at a group level is mild (mild functional disability HAQ-DI

score less than 1.0) (Khanna & Khanna, 2012), with the majority of observational studies reporting mean HAQ-DI scores of 1.0 or below (Alvarez-Hernandez et al 2008; Alvarez-Hernandez et al 2009; Becker et al 2009; Chandratre et al 2018; Fu et al 2017; Lopez Lopez et al 2017; Spaetgens et al 2015; Scire et al 2013; ten Klooster et al 2011; ten Klooster et al 2014; van Groen et al 2010; Wallace et al 2016). In addition, lower levels of disability were reported in people with gout in comparison with other diseases such as rheumatoid arthritis (Van Groen et al 2010).

The disease-specific GIS places a stronger emphasis on emotional consequences of gout, in comparison to SF-36 and HAQ-DI (Spaetgens et al 2014). Patient-rated gout severity has been shown to correlate with GIS scores in both cross-sectional and prospective studies (Hirsch et al 2008; Khanna et al 2011b), with participants with more severe gout having higher GIS scores indicating a greater impact of gout on HRQOL and worse disease-specific HRQOL. The five subscales of the gout impact scale cover different domains of disease-specific HRQOL; gout concern overall (CO), medication side effects (MSE), unmet treatment need (UTN), wellbeing during an attack (WBDA), and concern during an attack (CDA).

1.9.3 Change in HRQOL in people living with gout over time

The majority of studies investigating HRQOL in people living with gout are cross-sectional (see appendix 1), thus fail to capture the dynamic nature of HRQOL. Of the prospective studies which report HRQOL in people with gout, only a small number explicitly describe the change in HRQOL measures over time by describing mean change, effect sizes, and standardised response means (SRM).

Gout impact scale (GIS)

Mean change (SD) ranging from -2.0 (22.6) to -9.1 (21.2) were observed in different GIS subscales between baseline and one year, indicating an improvement in HRQOL, in an

American observational cohort study (Wallace et al 2016). Standardised response means (SRM) ranging from 0.09 to 0.43 for the different GIS subscales, and effect sizes ranging from 0.07 to 0.34 for different GIS subscales were observed for the same cohort (Wallace et al 2016).

Mean change (SD) in different GIS subscales ranging from -5.24 (14.26) to -7.61(18.76) were observed between week four and week eight in a subgroup of RCT participants deemed to have 'minimally improved' gout based on specific anchor questions (Khanna et al 2011b). Effect sizes for different GIS subscales ranging from 0.22 to 0.38 were observed for the same group of participants (Khanna et al 2011b).

More recently higher standardised response means (SRM) for GIS subscales over two years, ranging from 0.18 to 1.08, were reported by Chinchilla, Doherty & Abhishek (2019) for participants whose gout status was classified as 'better' after two years in an RCT. This group of participants were recruited from GP practices in the UK and included both participants receiving the nurse-led urate-lowering intervention and those receiving usual care.

SF-36 PF10

An effect size of 0.48 was observed in SF-36 Physical function (PF10) scores from baseline to one year, displaying an improvement in HRQOL, in a Spanish prospective cohort with participants who attended specialist gout clinics (Khanna et al 2011a). Improvement in HRQOL was also reported between baseline and 25 weeks in an RCT by Strand et al (2012), with mean change (SD) in SF-36 Physical function (PF10) scores of 0.25 (18.97) and 11.8 (24.1) in the placebo and intervention group respectively.

HAQ-DI

Change in HAQ-DI scores over six months was reported in a Mexican prospective study of participants from rheumatology departments by Alvarez-Hernandez et al (2008). No change

in HAQ-DI score was seen in 28.7% of participants, 38.3% experienced an improvement, and 32.9% experienced a deterioration. An effect size of 0.62 was reported for HAQ-DI scores from baseline to six months in this cohort.

The studies which have investigated HRQOL prospectively predominantly include participants from secondary care and often involve a clinical intervention. Studies in primary care often combined participants with those from rheumatology clinics, thus not permitting separate interpretation of HRQOL in people in primary care (Dalbeth et al 2013; Singh et al 2016; Stewart et al 2018; Wallace et al 2016). Hence results may not be generalizable to primary care populations where most patients with gout are managed. The majority of studies prospectively investigating HRQOL follow people with gout for one year or less, thus only permitting a short period of time to capture any change in HRQOL.

Thus, there is a paucity of observational studies prospectively investigating change in both disease-specific and generic HRQOL in people with gout, specifically in primary care and over periods longer than one year.

1.9.4 Factors associated with HRQOL in people living with gout

A range of gout-specific, comorbid, socio-demographic and other factors have been reported to be associated with HRQOL in people living with gout. A summary of different factors identified as associated with either generic or gout-specific HRQOL can be found in appendix 2. The key findings of these cross-sectional and prospective studies, relating to the factors associated with HRQOL, are described here.

Gout characteristics

Flare frequency

Gout flares have been highlighted as the major factor associated with HRQOL in people with gout (Chandratre et al 2013) with flare frequency associated with worse disease-specific

HRQOL on some of the GIS subscales (Chandratre et al 2018; Hirsch et al 2008), worse generic HRQOL in the physical domain (Becker et al 2009; Khanna et al 2011a; Khanna et al 2012c; Proudman et al 2019; Scire et al 2013) and greater activity limitation measured via HAQ-DI scores (Scire et al 2013). In prospective studies flare frequency has been associated with change in both SF-36 PF10 scores (Khanna et al (2011a) and HAQ-II scores (Stewart et al 2018) with more frequent flares associated with worse scores.

Current/recent flare

Experiencing a current or recent flare has been associated with worse GIS subscale scores (Chandratre et al 2018; Khanna et al 2015; Lee et al 2019), worse generic HRQOL measured via SF-12 PCS and SF-36 (Khanna et al 2015; Scire et al 2013) and worse HAQ-DI scores (Chandratre et al 2018; Scire et al 2013). In addition, a flare within the previous three months was associated with a worsening total GIS score (summary of all GIS subscales) over one year in a prospective study by Wallace et al (2016).

Number of joints affected

Several studies have reported worse gout-specific HRQOL (Chandratre et al 2018; Hirch et al 2010; Wallace et al 2016), generic HRQOL on the physical or mental domain (Becker et al 2009; Khanna et al 2011a; Lee et al 2009; Scire et al 2013; Wallace et al 2016) or activity limitation (Alvarez-Hernadez et al 2008; Becker et al 2009; Chandratre et al 2018; Scire et al 2013) where more joints are affected by gout. Tender joint count was associated with change in activity limitation, displayed by worse HAQ-II scores, at one year in a prospective analysis by Stewart et al (2018).

Gout severity/activity

Patient or physician rated 'gout severity' or 'gout activity' have correlated with worse HRQOL measured via the GIS subscales (Chincilla, Doherty & Abhishek, 2019; Hirsch et al 2008; Hirsch

et al 2010; La-Crette et al 2018; Sarkin et al 2010) or measured via HAQ-DI (Wallace et al 2016). Wallace et al (2016) also reported that a worsening physician severity score was associated with worse prospective total GIS scores.

Disease duration, serum urate levels and medications to treat gout

Studies investigating the relationship between HRQOL and other gout-specific factors such as disease duration, tophi, serum urate levels, and medication to treat gout, have revealed conflicting findings. Living with gout for longer has been associated with better disease-specific HRQOL measured via GIS unmet treatment need (UTN) subscale (Chandratre et al 2018) but worse generic HRQOL measured via HAQ-DI and SF-36 MCS (Scire et al 2013). However other studies have reported no association between disease duration and SF-36 or HAQ-DI scores (Becker et al 2009; Chandratre et al 2018). In addition, Wallace et al (2016) eliminated disease duration, along with tophi, and serum urate level, as covariates from multivariable analysis as they were not associated with change in total GIS score.

The presence of tophi has been associated with worse scores for GIS concern overall (CO) (Hirsch et al 2010), SF-36 PCS (Becker et al 2009; Scire et al 2013), SF-36 PF10 (Becker et al 2009; Edwards et al 2019), SF-12v2 (Khanna et al 2012), and HAQ-DI (Alvarez-Hernadez et al 2008; Edwards et al 2019; Scire et al 2013). However, no association between a history of tophi and any of the GIS subscales (Chandratre et al 2018; Lee et al 2019) or HAQ-DI scores (Becker et al 2009) has been reported.

Elevated serum urate levels have been associated with worse disease-specific HRQOL measured via GIS concern overall (CO) and GIS unmet treatment need (UTN) (Hirsch et al 2008; Lee et al 2019), generic HRQOL on the physical domain measured via SF-36 PCS (Khanna et al 2011a) and HAQ-II scores (Stewart et al 2018). However, only a minimal correlation between the latest serum urate level and GIS concern overall was observed by Hirsch et al

(2010), and no correlation between serum urate and HAQ-DI or any of the SF-36 subscales was observed by Becker et al (2009). In addition, no difference in generic HRQOL, measured via WHO-QOL BREF, was observed in participants dichotomized by serum urate levels (Roddy, Zhang & Doherty, 2007c). Serum urate was not associated with change in HAQ-II scores over time (Stewart et al (2018).

Several observational studies have reported no association between allopurinol use and generic HRQOL measured by SF-36, HAQ-DI (Chandratre et al 2018; Scire et al 2013), SF-12 (Proudman et al 2019) or WHO-QOL BREF (Roddy, Zhang & Doherty, 2007c). However, Chandratre et al (2018) reported that using allopurinol was associated with better disease-specific HRQOL measured via GIS unmet treatment need (UTN) and with worse disease-specific HRQOL measured via GIS wellbeing during an attack (WBDA). Worse SF-36 PCS and HAQ-DI scores have been associated with colchicine (Fu et al 2017; Scire et al 2013) and NSAID use (Scire et al 2013) in other observational studies, whereas Alvarez-Nemegyei et al (2005) described how NSAID requirement was not associated with disability, indicated by a HAQ-DI score greater than 0.5. However, caution is required when interpreting the findings of observational studies and associations between use of medication such as ULT and outcomes should not be interpreted as causal effects.

Comorbid characteristics

Number and type of comorbidities

A number of studies have reported associations between generic HRQOL and comorbidities. Worse SF-36 PCS and HAQ-DI scores have been associated with experiencing one or more comorbidities (Becker et al 2009), the total number of comorbidities (Scire et al 2013; Spaetgens et al 2015), and the Charlson Comorbidity Index (Spaetgens et al 2015). Medical and musculoskeletal comorbidities were associated with reduced generic HRQOL overall, measured via WHO-QOL BREF, across all domains except the social domain (Roddy, Zhang &

Doherty, 2007c). Alvarez-Nemegyei et al (2005) reported that hyperlipidaemia and ischaemic heart disease (IHD) were associated with disability measured via the HAQ-DI. A range of comorbidities including renal disease, diabetes, and hypertension were correlated with both worse SF-36 PCS and MCS, indicating worse generic HRQOL in the study by Lee et al (2009). Whilst various comorbidities including renal disease, diabetes, angina, and stroke, were associated with worse generic HRQOL, measured via both SF-36 PF10 and HAQ-DI (Chandratre et al 2018). Type II diabetes was associated with a worse HAQ-II score after one year but only for participants who experienced no activity limitation (had a HAQ-II of 0) at baseline (Stewart et al 2018).

However, studies investigating the relationship between comorbidities and disease-specific HRQOL have yielded less consistent findings. Lee et al (2019) found that participants with more comorbidities had better disease-specific HRQOL measured via the GIS concern overall (CO) and GIS medication side effects (MSE), compared to those participants with less comorbidities. Apart from an association with worse GIS concern overall score and hypertension, no associations between individual self-reported comorbidities and any of the GIS subscales was reported by Chandratre et al (2018). No association between either GIS subscales or total GIS score and individual comorbidities (Wallace et al 2016), the number of comorbidities (Spaetgens et al 2015), or the Charlson Comorbidity Index (Spaetgens et al 2015; Wallace et al 2016) have been observed.

Anxiety and depression

Both worse disease-specific and generic HRQOL has been observed in participants reporting anxiety or depression. Participants who indicated some or severe problems on the anxiety and depression items of the EQ-5D-3L questionnaire in a study by Lee et al (2019) had higher scores on all of the GIS subscales, compared with those who did not report problems relating to anxiety and depression. Worse disease-specific HRQOL, indicated by higher scores on all of

the GIS subscales, and worse generic HRQOL, indicated by SF-36 PF10 and HAQ-DI scores, has been associated with both depression and anxiety (Chandratre et al 2018). Conversely Fu et al (2017) reported that whilst depression was associated with both worse generic HRQOL in the physical domain, measured via SF-36 PCS and HAQ-DI, anxiety was only associated with worse generic HRQOL on the mental domain, measured using SF-36 MCS.

Pain

The relationship between pain and HRQOL in people living with gout has been investigated by a range of studies. Worse GIS subscale scores have been associated with more time between gout flares with gout-related pain (Hirsch et al 2010), typical gout pain measured via an 11-point visual analogue scale (VAS) (Hirsh et al 2010) and the presence of body pain (Chandratre et al 2018). Worse SF-36 PCS, SF-36 PF10, SF-36 MCS, HAQ-DI or HAQ-II scores have been associated with the presence of body pain (Chandratre et al 2018), worse gout-related pain (Becker et al 2009), more time between flares with pain (Lee et al 2009), and greater pain severity on a Likert scale or visual analogue scale (VAS) (Fu et al 2017; Lee et al 2009; Lopez Lopez et al 2017; Stewart et al 2018). VAS score was also associated with change in HAQ-II scores at one year, but this was only in participants who had a HAQ-II score of 0 at baseline (Stewart et al 2018).

Socio-demographic characteristics

Age

Studies investigating the relationship between age and HRQOL in people living with gout have indicated that older age is associated with worse generic HRQOL on the physical domain but better disease-specific HRQOL and generic HRQOL in the mental domain. Older age has been associated with increased HRQOL on all the WHO-QOL domains, except for HRQOL on the physical domain where older age was associated with worse HRQOL (Roddy, Zhang & Doherty

2007c). Older age was also associated with better scores for SF-36 MCS (Lee et al 2009; Wallace et al 2016) but worse SF-36 PCS scores (Lee et al 2009; Khanna et al 2011), SF-36 PF10, HAQ-DI (Chandratre et al 2018) and HAQ-II scores (Stewart et al 2018). Older age was associated with worse HAQ-II scores at one year in a prospective study by Dalbeth et al (2011). Wallace et al (2016) found that older age was associated with an improved disease specific HRQOL, measured by total GIS score, in people living with gout over one year. Better disease-specific HRQOL with older age was also reported on all the GIS subscales, except GIS unmet treatment need (UTN), by Chandratre et al (2018).

Sex

Few studies have identified associations between sex and HRQOL in people living with gout. Roddy, Zhang & Doherty (2007c) reported that being male was associated with a better HRQOL on the psychological domain (measured via WHO-QOL BREF), but a reduced HRQOL in the social domain in people living with gout. In contrast Chandratre et al (2018) reported that being female was associated with worse generic HRQOL measured by both SF-36 PF10 and HAQ-DI.

Ethnicity

Chandratre et al (2018) observed an association between being Caucasian and better GIS concern during an attack (CDA) and medication side effects (MSE) scores but no association with SF-36 PF10 or HAQ-DI scores. Whilst Singh et al (2016) described worse disease-specific HRQOL, demonstrated by higher GIS unmet treatment need (UTN), wellbeing during an attack (WBDA), and concern during an attack (CDA), in African American participants with gout in comparison with Caucasian participants. African American participants also had worse generic HRQOL on the mental domain, indicated by lower SF-36 MCS score, however there was no difference in HAQ-DI scores between the two groups (Singh et al 2016). In another prospective study of people living with gout Dalbeth et al (2013) identified worse HAQ-II and

SF-36 PF10 scores, in Maori and Pacific participants in comparison with non-Maori or non-Pacific participants.

Education

Attendance at further education was associated with better generic HRQOL, measured via SF-36 PF10 in the study by Chandratre et al (2018), however no relationship between attendance at further education and either HAQ-DI scores or GIS scores was reported in this cohort. Educational attainment was not associated with higher HAQ-DI scores in a multi variable analysis by Alvarez-Nemegyei et al (2005). Conversely, educational attainment to upper secondary level or beyond has been associated with better generic HRQOL and activity limitation, measured via SF-36 PCS and HAQ-DI respectively (Scire et al 2013).

Other socio-demographic factors

Chandratre et al (2018) also investigated the relationship between other socio-demographic factors and HRQOL in people living with gout. Being classified as most deprived based on indices of multiple deprivation (IMD) was associated with worse SF-36 PF10 scores, but not worse GIS scores. Whilst not being married or cohabiting was associated with worse SF-36 PF10 scores, it was not associated with either worse HAQ-DI scores or GIS scores.

BMI and alcohol frequency

BMI

Greater BMI was associated with a lower HRQOL on all domains except the social domain in the multivariable analysis by Roddy, Zhang & Doherty (2007c). Obesity or a BMI ≥35 kg/m^2 has been associated with worse SF-36 PCS (Scire et al 2013), SF-36 PF10 (Chandratre et al 2018) and HAQ-DI scores (Chandratre et al 2018; Scire et al 2013), but not GIS subscales scores (Chandratre et al 2018). Stewart et al (2018) identified an association between a higher BMI and worse HAQ-II at baseline, but not change in HAQ-II at one year.

Alcohol frequency

Compared to consuming alcohol daily, consuming alcohol only on special occasions or never consuming alcohol was associated with worse SF-36 PF10 and HAQ-DI scores, and consuming alcohol only on special occasions was associated with worse GIS unmet treatment need (UTN) scores.

In summary, disease-specific and generic HRQOL in people living with gout has been associated with several gout-specific, comorbid, socio-demographic and other factors in previous studies. Such factors include gout flares, experiencing a current or recent flare, the number of joints affected by gout, disease duration, serum urate level, medication used to treat gout, number or type of comorbidities, anxiety, depression, pain, age, sex, ethnicity, BMI, educational attainment and other factors. Of all of the factors identified gout flares are the major factor associated with impaired disease-specific and generic HRQOL in people living with gout. However, there is a paucity of studies which have investigated the factors associated with HRQOL, whilst adjusting for a range of gout-specific, comorbid, socio-demographic and other factors. The majority of studies which investigated factors associated with HRQOL in people living with gout only report associations observed in cross-sectional analyses, with a paucity of studies investigating factors associated with change in HRQOL particularly in primary care. The studies which have investigated factors associated with change in HRQOL in people with gout have done so over periods no longer than 12 months. This paucity in the literature justifies the investigation undertaken within the following chapters.

1.10 Conclusion

In conclusion, this chapter provides background and context to the topics which are covered in this thesis. This chapter highlights the paucity of research relating to gout flare trajectories, change in HRQOL over time, and factors associated with change in HRQOL in people living with

gout in primary care. In the following chapter the structure of this thesis, the overall aim of each chapter and the key objectives of the thesis are outlined.

2 Chapter Two Overview of thesis

2.1 Overview of chapter and aim

The previous chapter provided relevant background to gout, gout flares, and health-related quality of life. The aim of the following chapter is to outline the structure of the thesis, the overall aim of each chapter and the key objectives of the thesis.

2.2 Thesis statement

Gout is a common inflammatory arthritis which, despite the availability of effective treatment, is often sub-optimally managed in primary care (Kuo et al 2015a). The clinical hallmark of gout is rapid onset extremely painful acute inflammatory flares, characterised by swelling, erythema, heat and joint tenderness (Roddy, Mallen & Doherty, 2013; Taylor et al 2015; Teng, Nair & Saag, 2006). Gout is associated with impaired HRQOL, and the impact on HRQOL is greater in the physical domain (Chandratre et al 2013). Gout flares are the major factor associated with impaired HRQOL in people living with gout but further factors associated with poor HRQOL include other gout-specific, comorbid, socio-demographic and other factors (Chandratre et al 2013).

Little is known about change in gout flares or HRQOL over time in people living with gout in primary care.

2.3 Structure of thesis chapters and aim of each chapter

Chapter One	**Introduction and background**
Aim	To introduce the thesis and provide relevant background to the topic.
Chapter Two	**Overview of thesis**
Aim	To outline the structure of this thesis, the overall aim of each chapter and the key objectives of the thesis.
Chapter Three	**Prospective cohort study**
Aim	To describe the prospective cohort study in which this thesis is nested.
Chapter Four	**Description of cohort and assessment of bias**
Aim	To describe the flow of participants through this cohort study and investigate attrition bias and missing data.
Chapter Five	**Frequency and pattern of gout flares**
Aim	To describe the frequency and pattern of self-reported gout flares in a prospective cohort of people living with gout in primary care.
Chapter Six	**Gout flare trajectory classes**
Aim	To identify and describe latent classes of distinct gout flare trajectories in a prospective cohort of people living with gout in primary care.
Chapter Seven	**Change in health-related quality of life**
Aim	To describe how gout-specific and generic HRQOL in people living with gout changes over a three-year period
Chapter Eight	**Factors associated with change in health-related quality of life**
Aim	To describe the factors associated with change in gout-specific and generic HRQOL in people living with gout in primary care over a three-year period.
Chapter Nine	**Discussion and conclusion**
Aim	To discuss the key findings of the thesis, considering the strengths, limitations, implications for clinical practice, implications for future research, and to draw conclusions.

Figure 2.1 Title and overall aim of each thesis chapter

2.4 Key thesis objectives

Chapter 4 objectives

1. Describe the response and the flow of participants through this cohort study.
2. Compare characteristics of responders and non-participants at the five questionnaire mail-out time points (i.e. Baseline, 6, 12, 24, 34 months).
3. Compare characteristics of participants who responded on a different number of occasions (i.e. only baseline, baseline plus one other time-point, baseline plus two other time-points, baseline plus three other time-points, all time-points).
4. Determine the number and percentage of missing item responses of key variables within the questionnaire and medical record review data; indicating the type and volume of missing item response data.

Chapter 5 Objectives

1. Use descriptive statistics and spaghetti plots to describe the frequency and pattern of self-reported gout flares.
2. Use latent growth curve modelling (LGCM) to determine the shape of the growth curve which best describes the longitudinal gout flare data.

Chapter 6 Objectives

1. Use latent class growth analysis (LCGA) and growth mixture modelling (GMM) to identify distinct latent classes of gout flare trajectories in people living with gout in primary care.
2. Describe the characteristics of members of each gout flare trajectory class.

Chapter 7 Objectives

1. Describe change in GIS subscales (CO, MSE, UTN, WBDA, CDA), SF-36 PF10 and HAQ-DI scores over a three-year period.
2. Describe the correlation, and distribution of the GIS subscales (CO, MSE, UTN, WBDA, CDA), SF-36 PF10 and HAQ-DI scores over a three-year period.
3. Describe the GIS subscales, SF-36 PF10 and HAQ-DI scores over three years for participants in each distinct latent gout flare trajectory class identified in chapter 6.

Chapter 8 Objective

1. Use linear mixed modelling to describe the variables associated with change in GIS subscales (CO, MSE, UTN, WBDA, CDA), SF-36 PF10, and HAQ-DI scores over a three-year period.

3 Chapter Three Prospective Cohort Study

3.1 Overview of chapter and aim

The previous chapter outlined the structure of the whole thesis and described the aims and key objectives of this thesis. The aim of the following chapter is to describe the prospective cohort study in which this thesis is nested. The aim of the study was to investigate health-related quality of life (HRQOL) in a large cohort of people living with gout in primary care (Chandratre et al 2012).

3.2 Study design

The study was a prospective cohort study utilising a postal questionnaire survey of people with gout based in primary care. Figure 3.1 displays the seven phases of the cohort study which include the five time-points where participants were mailed a postal questionnaire. The cohort study was established as part of a previous PhD studentship and the study protocol has been described and published previously (Chandratre et al 2012). In addition to a postal questionnaire, the study included an electronic review of medical records and qualitative focus group interviews. The baseline findings from this prospective cohort study (Chandratre et al 2018) and nested qualitative study (Chandratre et al 2016) have been published previously. This thesis used data derived from baseline and the postal follow-up of the cohort over a three-year period and medical record review. The thesis used data from the entire cohort but only a selected number of outcomes from the study questionnaire were analysed as part of this thesis. The author's role in relation to the cohort study included writing data analysis plans, preparing data for analysis (e.g. building databases, cleaning data, recoding and categorising variables), running PPIE group meetings, undertaking statistical analysis, and reporting and communicating findings.

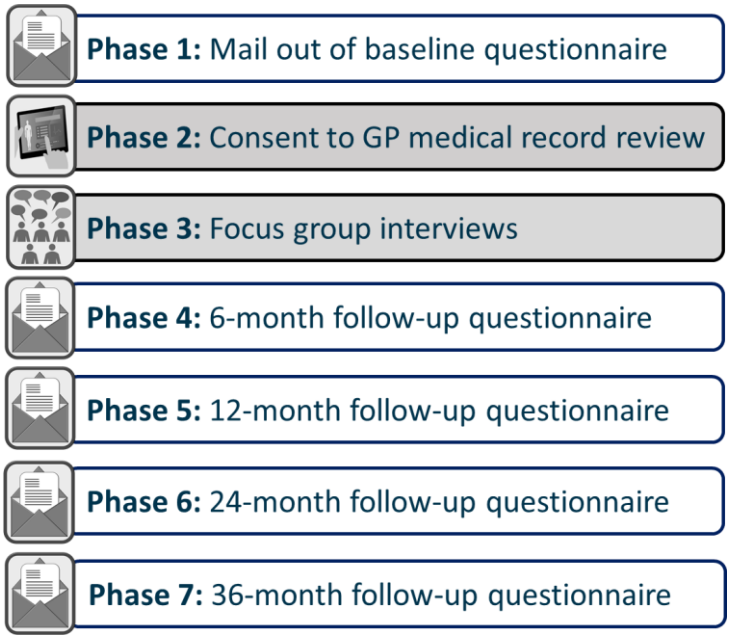


Figure 3.1 Phases of prospective cohort study

3.3 Participants and setting

Study population

All adults (over the age of 18 years) registered with 20 general practices within the West Midlands who had either:

- Consulted primary care in the preceding two years for gout or,
- Received a prescription for either colchicine or allopurinol in the last two years

Inclusion criteria

- Registration with a participating general practice during the course of the study
- Read code consultation for gout or prescription of colchicine or allopurinol during the preceding two years
- Provided written informed consent to participate in the study

Exclusion criteria

- Under the age of 18 years of age
- Vulnerable groups e.g. significant dementia, severe enduring mental illness, active cancer or other terminal illness.

Identification of participants

Eligible participants for the study were identified via general practice Read codes assigned to consultations or prescriptions relating to gout. Read codes are alphanumeric medical codes which allow primary care health care workers to code clinical consultations (Chisolm, 1990; Porcheret et al 2004) and have been shown to be a valid indicator of the true presence of disease; with Read codes for gout displaying a sensitivity of 93.7% and positive predictive value (PPV) of 97.6% (Hassey, Gerret & Wilson, 2001). Whilst a diagnosis of gout in CPRD has been validated with a PPV of 90% reported (Meier & Jick, 1997). Table 3.1 details the Read codes used by the Primary Care Centre Versus Arthritis (PCCVA) at Keele University for participant identification. Clinical staff in the local area have received training on the use of Read codes and data entry as part of a programme of assessment, feedback and training (Porcheret et al 2004).

Table 3.1 Read codes used by the Primary Care Centre Versus Arthritis

Code	Term
C34	Gout
N023	Gouty Arthritis
EGTON 227	Gout NOS
OX2740G	Gout Acute/ ox
1443	H/O: gout
EMISR4QG01	Gouty tophi + Gout NOS
2D52	O/E-auricle of ear-tophi
669	Gout Monitoring

The GP practices have a formal agreement with the West Midlands Clinical Research Network

(CRN) and staff contracted to work for the practice undertake various functions including the identification of patients for the participation in research. Staff from the West Midlands CRN health informatics team ran a one-off electronic search of primary care records from participating practices in order to identify patients who in the preceding two years had had a consultation for gout or a prescription for colchicine or allopurinol. The lead general practitioner at each general practice was requested to identify vulnerable patients from the list obtained who would be excluded from the study. To ensure patients were not contacted inappropriately, the CRN team screened the mailing list prior to the first and subsequent mailing to identify patient deaths or departures from the practice.

Contacting participants

All eligible patients were mailed a study pack from their GP which was presented on the GP practice headed notepaper. The baseline mail-out took place in October & November 2012. The study pack included:

- A cover letter inviting patients to take part in the study (see appendix 3)
- A participant information sheet (PIS) (see appendix 4)
- A pre-paid envelope for the return of the questionnaire and consent form
- Baseline questionnaire and attached consent form (see appendix 5)

Reminder postcards (see appendix 6) were sent out if there was no response after two weeks and a repeat questionnaire with PIS and covering letter (see appendix 7) was mailed if there was no response after four weeks.

Follow-up out to participants

For each follow-up mail-out at 6, 12, 24 and 36 months the following process was followed; removal of deaths and departures from the mailing database, mail-out of questionnaire with covering letter, reminder postcard sent if no response in two weeks and repeat questionnaire was mailed if there was no response in four weeks.

Medical note review

Individuals who consented to medical record review had their computerised medical records tagged by a member of the CRN. A medical record review included a review of records over the two years prior to recruitment and the three year follow-up period to collect anonymised consultation data relating to diagnosis, investigations, treatment and comorbidities. Specialist health informatics staff linked to GP practices undertook the medical record review.

Patient Public Involvement and Engagement

Patient public involvement and engagement (PPIE) was incorporated in the design and management stages of the study. Advertisements, via posters and flyers, were used to find patients with a diagnosis of gout to take part in the PPIE for this study. Lay members of the PPIE group provided feedback on conceptual and design aspects of the questionnaire survey. Patient engagement has been proposed to improve participant enrolment and decrease attrition (Domecq et al 2014; Shippee et al 2013). PPIE members were involved in interpretation of findings and planning dissemination of results.

3.4 Ethics

Approval

Ethics approval for the study was granted by National Research Ethics Service (NRES) committee North West Liverpool East Research and Ethics Committee (REC) reference number 12/NW/0297. A copy of the ethical approval letter can be found in appendix 8.

Consent

The consent form was included in the study pack attached to the baseline questionnaire. Participants were asked to consent separately to medical record review and future postal contact.

Confidentiality

Each participant was allocated a unique study identification number. Research data obtained from responses to the study questionnaire were recorded on a confidential and anonymised study data entry database which is separate from personal mailing data, so that research data is not linked to personal data.

3.5 Data management

Data input

For quality assurance, one in ten questionnaires were checked for accuracy of coding by a member of the study team, corrections were made to the data entry database as necessary. Comments and queries which arose whilst returned questionnaires were inputted to the data entry database were recorded on a comments and queries spreadsheet along with any actions taken or lessons learnt which could be applied to any future queries with data entry.

Data cleaning

The data were checked using descriptive statistics to detect and then correct any potential errors (Pallant, 2016; Van de Broeck et al 2005; Osborne, 2013). Data cleaning is advocated to promote data standardisation and data quality (Van de Broeck et al 2005). This inspection and correction of error within data sets can help to prevent later difficulties with the analysis and interpretation of data (Pallant, 2016; Osborne, 2013).

Sample size

A sample size calculation was undertaken at the inception of this study. Over 1,800 people were contacted at baseline (Chandratre et al 2012). The sample size calculation was based on the assumption of a 70% response rate at baseline and a 30% drop out over the three years, and indicated that in order to use the information recorded at all five time-points a sample of

882 would allow a smallest meaningful difference in HRQOL of 0.2 standard deviation units to be detected between two groups (441 participants per group) defined in terms of frequency of gout attacks (<2 attacks, ≥2 attacks per year) using a linear mixed model (significance 0.05, power 90%, autocorrelation 0.8) (Chandratre et al 2012).

3.6 Questionnaire content

A summary of the content of the questionnaire mailed to participants can be found in Table 3.2. The baseline questionnaire contained seven sections (see appendix 5 for a copy of baseline questionnaire and appendix 9 for tables of the content of each questionnaire section). Sections 3.6.1 to 3.6.5 on the following pages describe the questionnaire content used in the analysis in this thesis.

Table 3.2 Summary of questionnaire content at five mail out time-points

Questionnaire content	Baseline	6 months	12 months	24 months	36 months
Gout					
Frequency of gout flares	✓	✓	✓	✓	✓
Age at gout diagnosis	✓				
Occurrence of current gout flare	✓	✓	✓	✓	✓
History of oligo-/ polyarticular gout	✓	✓	✓	✓	✓
Use of allopurinol	✓	✓	✓	✓	✓
How gout affects life					
GIS	✓	✓	✓	✓	✓
IPQ-R	✓	✓	✓	✓	✓
General health					
SF-36 PF 10	✓	✓	✓	✓	✓
HAQ-DI	✓	✓	✓	✓	✓
Pain NRS	✓	✓	✓	✓	✓
Global assessment of health	✓	✓	✓	✓	✓
Comorbidities	✓				
How participants feel					
PHQ-9	✓		✓		✓
GAD-7	✓		✓		✓
Foot & other joint problems					
Hallux Valgus	✓		✓		✓
Foot pain, aching & stiffness	✓		✓		✓
Foot pain location	✓		✓		✓
Foot function	✓		✓		✓
Health care consultations regarding feet	✓		✓		✓
Body pain	✓		✓		✓
How gout affects work					
Current employment status	✓		✓		✓
Taken time off work due to gout in last 6 months	✓		✓		✓
Ability to undertake usual employment	✓		✓		✓
Inability to undertake normal employment attributed to joint problems	✓		✓		✓
Demographics & socioeconomics					
Date of birth	✓				
Sex	✓				
Relationship status	✓				
Education	✓				
Ethnic origin	✓				
Height, weight & alcohol					
Height	✓	✓	✓	✓	✓
Weight	✓	✓	✓	✓	✓
Alcohol consumption	✓				

GIS, Gout Impact Scale; **IPQ-R**, Revised Illness Perception Questionnaire; **SF-36 PF 10,** 36-Item Short-Form Health Survey 10-item Physical Function subscale**; HAQ-DI,** Health Assessment Questionnaire Disability Index**; NRS,** numerical rating scale**; GAD-7,** Generalised Anxiety Disorder-7**; PHQ-9,** Patient Health Questionnaire-9.

3.6.1 Questions about gout

Participants were asked to report how many flares of gout they had experienced in the last 12 months, or in the last six months at the six and 12-month time-point, using six response options (0, 1, 2, 3, 4, ≥5). Self-reported gout flares have been compared to the criterion of an assessment by a rheumatologist and shown to have a sensitivity of 91% and negative predictive value of 96% (Gaffo et al 2012). Respondents were asked whether they were experiencing a flare of gout at the point of completion of the questionnaire and if they had ever had gout affecting more than one joint (a history of oligo-/polyarticular flares). Participants were asked their age at diagnosis of gout. Gout disease duration was calculated by subtracting a participant's age at gout diagnosis from their current age at baseline. Respondents were asked to state whether they were currently taking allopurinol. If allopurinol was taken, participants were asked to either select the dose taken (50 mg, 100 mg, 200 mg, 300 mg, 400 mg, 500 mg, 600 mg, 700 mg, 800 mg, 900 mg or 'don't know') or state, via free text, an alternative dose to the categories stated. Self-reporting of current medication use, collected via mail in a large cohort, has displayed acceptable agreement with pharmacy records (Drieling et al 2016).

3.6.2 Questions about how gout affects life

Gout Impact Scale (GIS)

The self-reported impact of gout on study participants was assessed using the disease-specific HRQOL measure, the Gout Impact Scale (GIS) (Hirsch et al 2008). The GIS originally formed the Gout Impact section of the Gout Assessment Questionnaire (GAQ) 2.0 (Hirsch et al 2008) and was developed from the earlier Gout Assessment Questionnaire (GAQ 1.0) (Colwell et al 2006). The GIS consists of 5 subscales; gout concern overall (GIS CO) (first four items), gout medication side effects (GIS MSE) (two items), gout unmet treatment need (GIS UTN) (three

items), wellbeing during an attack (GIS WBDA) (11 items) and gout concern during an attack (GIS CDA) (four items) (Hirsch et al 2008). This totals 24 items each with response options provided on a five-point ordinal scale (Hirsch et al 2008). Each scale is scored from 0-100 with a higher score indicating worse HRQOL (Khanna et al 2011b). Changes in the GIS subscales of 5 to 8 points have been estimated as minimal clinically important differences (MCID) in patients with gout, however MCID results were not reported for the GIS MSE subscale (Khanna et al 2011b). See Table 3.3 for details relating to recall period, calculation and interpretation of scores.

Content validity has been demonstrated for the GIS and GAQ 1.0 by the involvement of rheumatologists and patient groups in the development of the scale (Colwell et al 2006; Hirch et al 2008). Patient-rated gout severity has been shown to correlate with all GIS subscale scores demonstrating construct validity (Hirsch et al 2008). Further construct validity has been displayed by the correlation of total GIS scores with HAQ-DI and SF-36 scores, and higher GIS scores in groups of patients who had experienced recent gout attacks (Wallace et al 2016). In a longitudinal study of gout patients, the GIS has been shown to be responsive to change with effect sizes of 0.07 to 0.34 and standardised mean response (SMR) of 0.09 to 0.45 (Wallace et al 2016).

The GIS subscales have been shown to have good test-retest reliability (Hirsch et al 2008) and adequate internal consistency (Hirsch et al 2008; Khanna et al 2011b). However, lower Cronbach alpha scores have been reported for the two-item GIS MSE and the three-item GIS UTN (Wallace et al 2016). Readability scores for the GIS suggest that a reading age of 12 years or older is required (Janssen et al 2019).

Table 3.3 Recall period, calculation of score and interpretation of score for measures about how gout affects life, general health and how participants feel

Section of questionnaire	Measure	Recall period	Calculation of score	Interpretation of score	
How gout affects your life	**GIS subscales**	Not specified	Item response rescaled and reversed as necessary to score from 0-100. Score is the mean of the responses for each subscale.	Each GIS subscale scored from **0-100**. **Higher scores** on each scale indicating a greater impact of gout/ **worse HRQOL**.	
About general health	**SF-36 PF 10**	Current situation "health now"	Scores for all items are coded, summed and then transformed on to a scale of 0 to 100	SF-36 PF10 scored from **0-100**. **Lower scores** indicate **worse HRQOL**. **Low**= Limited a lot in performing all physical activities including bathing or dressing. **High**= Performs all types of physical activities including the most vigorous without limitations due to health.	
	HAQ-DI	Past week	Highest component score in each category determines score for category. Dependence on device, aid or help increases a lower score to 2. The 8 category scores are averaged in to an overall HAQ-DI score.	Overall HAQ-DI scored from **0 to 3**. **Higher scores** indicate **worse activity limitation**. **0** = No disability **3** = Complete disability	
	NRS Pain	Past one week	Score from 0 to 10	**0** = No pain **10** = Pain as bad as it can be	
	NRS Patient Global	Current situation	Score from 0 to 10	**0** = Very well **10** = Very poor health	
How participants feel	**PHQ-9**	Last two weeks	Addition of scores	**0** = Minimal depression **27** = Severe depression	**Minimal** 0-4 **Mild** 5-9 **Moderate** 10-14 **Moderately Severe** 15-19 **Severe** 20-27
	GAD-7	Last two weeks	Addition of scores	**0** = Minimal anxiety **21** = Severe anxiety	**Minimal** 0-4 **Mild** 5-9 **Moderate** 10-14 **Severe** 15-21

GIS, Gout Impact Scale; **SF-36 PF10,** 36-Item Short-Form Health Survey 10-item Physical Function subscale**; HAQ-DI,** Health Assessment Questionnaire Disability Index**; NRS,** numerical rating scale**; GAD-7,** Generalised Anxiety Disorder-7**; PHQ-9,** Patient Health Questionnaire-9.

3.6.3 Questions about general health

SF-36 PF10

Self-reported physical function was assessed in this study using the 10-item physical function (PF10) subscale of the medical outcomes study 36-Item Short-Form Health Survey (SF-36) (Ware & Sherbourne, 1992). The SF-36 is a generic measure of generic HRQOL and the SF-36 PF-10 consists of 10 questions relating to the ability to undertake different activities during a typical day, and responses are obtained via a three-point ordinal scale (Ware & Sherbourne, 1992). SF-36 PF10 is scored from 0-100 with a lower scale score suggesting worse HRQOL (Brazier et al 1992). A low score is interpreted as suggesting that an individual is limited a lot in performing all physical activities including bathing or dressing (Ware & Sherbourne, 1992; Ware et al 1995). A high score is interpreted as indicating that an individual is able to perform all types of physical activities including the most vigorous, without limitations due to health (Ware & Sherbourne, 1992; Ware et al 1995). A MCID in SF-36 PF10 scores in people with gout of 5 has been cited by Khanna et al (2011a). See Table 3.3 for details relating to the recall period, calculation and interpretation of scores.

A range of studies have demonstrated the construct validity of the SF-36 when used with patients with gout. Lower SF-36 scores were associated with the presence of tophi, co-morbidities and polyarticular disease in a prospective study by Khanna et al (2011a). SF-36 scores were correlated with HAQ-DI scores in both prospective (Becker et al 2009) and cross-sectional studies (ten klooster et al 2011). SF-36 PF10 scores have also been able to discriminate between different levels of general health (ten Klooster et al 2011). The SF-36 has been shown to be responsive to change in a longitudinal cohort of gout patients, with SF-36 PF10 demonstrating an effect size of 0.48 (Khanna et al 2011a). The SF-36 PF10 has demonstrated convergent validity through a strong association with serious medical

conditions and discriminant validity with patients with solely mental health conditions (Brazier et al 1992; McHorney, Ware & Raczeck, 1993; Ware et al 1995).

High internal consistency reliability of the SF-36 PF10, when used with people with gout, has been demonstrated by a Cronbach alpha score of 0.97 (ten Klooster et al 2011). Acceptable test-retest reliability of SF-36 subscales has been reported by Becker et al (2009) in a gout specific population. The SF-36 has been shown to have an acceptable internal consistency and test-retest reliability in a primary care based postal questionnaire survey in the UK (Brazier et al 1992).

The median readability of items within the SF-36 has been determined to be at a reading level of around 6th grade (Calderon et al 2006; Paz et al 2009). SF-36 has been endorsed as a patient reported outcome for the assessment of HRQOL by Outcome Measures in Rheumatology Meeting (OMERACT) in light of the acceptable validity, responsiveness to change and reliability (de Latour, Taylor & Dalbeth, 2015; Singh et al 2011a).

HAQ-DI

Self-reported activity limitation was assessed by the inclusion of the Health Assessment Questionnaire Disability Index (HAQ-DI) (Bruce & Fries, 2003). The HAQ-DI is a generic measure of physical disability, which constitutes dimension one of the full Health Assessment Questionnaire and is widely used in rheumatology research (Bruce & Fries, 2003). The HAQ-DI consists of 20 items, arranged in eight categories of functioning relating to dressing & grooming, arising, eating, walking, hygiene, reach, grip and activities (Bruce & fries, 2003). Respondents are asked to report how much difficulty they have experienced in the past one week whilst undertaking these tasks via a four-point ordinal response scale (ranging from 'without any difficulty' to 'unable to do') (Bruce & Fries, 2003). Respondents are also asked to indicate whether they utilise any aids, devices or assistance (Bruce & Fries, 2003). Each

scale is scored from 0-3 with higher scale scores indicating activity limitation. A HAQ-DI score of zero suggests no disability in contrast to a score of three which suggests complete disability (Bruce & Fries, 2003). The MCID cited by Singh et al (2016) in HAQ-DI scores in people with gout was 0.22. See Table 3.3 for details relating to the recall period, calculation and interpretation of scores.

The face validity of the HAQ-DI when used with people living with gout has been described (Alvarez-Hernandez et al 2008), however floor effects, where 20.5 to 42.2% of participants have a HAQ-DI score of zero, have also been reported (Alvarez-Hernandez et al 2008; Taylor et al 2008). HAQ-DI scores have been shown to correlate well with task performance thus providing evidence of criterion validity (Bruce & Fries, 2003). The construct validity of the HAQ-DI when used with patients with gout, has been demonstrated by correlation with SF-36 PF10 scores and visual analogue scales (VAS) pain scores (Alvarez-Hernandez et al 2008; Taylor et al 2008), along with the ability to discriminate between different levels of health and disease characteristics (Alvarez-Hernandez et al 2008; Van Groen et al 2010). The HAQ-DI has been shown to be sensitive and responsive to change in a six-month prospective study of gout patients, with an effect size of 0.62 (Alvarez-Hernandez et al 2008).

The test-retest validity of the HAQ-DI in subjects with gout has been shown to be acceptable, with intraclass correlation coefficients (ICC) of 0.76 to 0.84 (Alvarez-Hernandez et al 2008; Becker et al 2009). Internal consistency of the HAQ-DI, when used to measure function in subjects with gout, has been reported as excellent with Cronbach alphas of 0.91 to 0.97 (Alvarez-Hernandez et al 2008; Taylor et al 2008; ten Klooster et al 2010).

Due to the acceptable validity, responsiveness and reliability demonstrated in gout studies HAQ-DI has been endorsed by the 2010 Outcome Measures in Rheumatology Meeting (OMERACT 10) as a patient-reported outcome for measurement of activity limitation in chronic gout (Singh et al 2011a).

NRS Pain and Patient Global

The VAS pain scale and patient global assessment VAS from dimension two of the full HAQ have been endorsed as patient-reported outcomes in OMERACT 10 (Singh et al 2011a) due to the acceptable validity and reliability of these measures in studies measuring outcomes of gout (Singh et al 2011a; Singh et al 2011b). There was agreement at OMERACT 10 that the VAS used in the HAQ could be converted to numerical rating scales (NRS) when measuring patient-reported outcomes. An 11-point NRS used to measure pain (where 0 indicates 'no pain' to 10 indicates 'pain as bad as it can be') and a patient global assessment NRS (where 0 indicates 'very well' and 10 indicates 'very poor health') were included within this questionnaire (see Table 3.3).

Body pain

Assessment of body pain in this survey was undertaken by asking participants whether they had experienced any pain, ache, discomfort or stiffness lasting longer than one day in any part of their body during the past 4 weeks.

Comorbidities

Participants were asked to indicate in the baseline study questionnaire whether they had ever been diagnosed with or treated for a range of different comorbidities (diabetes, hypertension (HT), hyperlipidaemia (HL), myocardial infarction (MI), angina, cerebrovascular accident (CVA), transient ischaemic attack (TIA), renal failure (RF), renal calculi). Self-reporting of comorbidities has been shown to have acceptable test-retest reliability and correlation with documented medical record comorbidities (Sangha et al 2003; Simpson et al 2004).

3.6.4 Questions about depression and anxiety

PHQ-9

Depression in this study was assessed using the Patient Health Questionnaire-9 (PHQ-9). The PHQ-9 is a validated measure of depression severity which was adapted from the longer patient health questionnaire (Kroenke, Spitzer & Williams, 2001). The self-administered questionnaire consists of a checklist of nine items and responses via a four-point ordinal scale (Kroenke, Spitzer & Williams, 2001). Respondents are asked to recall how often they have been bothered by these nine different issues in the last two weeks (Kroenke, Spitzer & Williams, 2001). Scores indicate the severity of depression and it is the recommended tool for the measurement of depressive symptoms by the Diagnostic and Statistical Manual of Mental Disorder; with scores of 0-4 indicating minimal depression and 20-27 severe depression (Kroenke, Spitzer & Williams, 2001; Moriarty et al 2015). See Table 3.3 for details relating to the recall period, calculation and interpretation of scores.

The PHQ-9 has displayed criterion and construct validity when used with in primary care and both the test-retest reliability and internal reliability have been found to be excellent (Kroenke, Spitzer & Williams, 2001).

The association of gout severity with depression in this cohort study at baseline has been published previously (Prior et al 2016).

GAD-7

Anxiety was assessed using the Generalised Anxiety Disorder-7 (GAD-7). The GAD-7 is a brief scale devised to identify probable cases of generalised anxiety disorder (Spitzer et al 2006). This self-reported tool was adapted from the 13 item GAD scale (Spitzer et al 2006). Respondents are asked to state, by the selection of one of 4 ordinal responses, how often in the last 2 weeks they have been bothered by the issues described in the 7 items (Spitzer et al

2006). Scores specify the degree of general anxiety disorder with 5 to 9 indicating mild generalised anxiety and 15 to 21 indicating severe generalised anxiety (Spitzer et al 2006). See Table 3.3 for details relating to the recall period, calculation and interpretation of scores.

The GAD-7 has demonstrated construct validity when used within primary care (Spitzer et al 2006). The internal consistency of the GAD-7 has been shown to be excellent and the test-retest validity good (Spitzer et al 2006).

3.6.5 Questions about socio-demographics, height, weight and alcohol consumption

Participants were asked their date of birth, sex, and whether they had attended further education. The selection of one response from six nominal category responses was used in the questionnaire to collect data relating to relationship status (married, widowed, co-habiting, divorced, separated or single) and ethnic origin (White UK/European, Asian, Afro Caribbean, African, Chinese, or other) respectively.

Participants were asked their weight and height in either metric or imperial units. Self-reported weight and height has been shown to be highly correlated with measured weight and height and is thus valid for detecting associations between weight and height with disease in epidemiological studies (Spencer et al 2002; McAdams, Van Dam & Hu, 2007). The self-reported weight and height was used to calculate body mass index (BMI) calculated as weight (kg)/height (m^2) (Keys et al 1972). BMI calculated from self-reported weight and height has been shown to be highly correlated with measured BMI (Spencer et al 2002; McAdams, Van Dam & Hu, 2007). BMI was classified into the categories of underweight, normal range, overweight and obese using classification criteria specified by the World Health Organisation (World Health Organisation, 2000).

Self-reported alcohol consumption was recorded using a frequency and quantity approach, which has been described as a generally reliable and valid method to investigate alcohol intake

in surveys (Del Boca & Darkes 2003; Grant et al 1995). Participants were asked how often alcoholic drinks were consumed using six ordinal categories (never, special occasions, 1 to 3 times a month, once or twice a week, 3 to 4 times a week, daily almost daily) and how many alcoholic drinks of a specific volume were consumed in an average week. Participant postal codes were used to derive indices of multiple deprivation (IMD) (Department for communities and Local Government, 2011).

3.7 Questionnaire Design and Administration; attempts to maximise response and minimise bias.

Several aspects of this study's method, regarding the administration of the questionnaire, have been advocated to optimise the response rates in postal questionnaire surveys. The questionnaires were accompanied with a letter personalised with each participant's name (Edwards et al 2009; Sinclair et al 2012; Sahlqvist et al 2011). This personalised letter was from a named doctor, displayed sponsorship from a local University and explained the importance of the questionnaire (Edwards et al 2009; Bowling, 2014). Non-responders were sent reminders, which has been shown to be an economical and effective method to increase survey response rates (Edwards et al 2009; Sahlqvist et al 2011; Shih & Fan, 2008). Non-responders were posted a reminder card 14 days after the original posting of the questionnaire (Salhlqvist et al 2011). At 14 days after the postcard reminder non-responders were sent a further reminder with a second copy of the questionnaire included (Edwards et al 2009).

The questionnaire started with questions relating specifically to the medical condition of interest, gout, rather than general health questions thus increasing the saliency of the questionnaire to the respondent which has been shown to increase response time and response rate (Dunn, Jordan & Croft, 2003). Specifically, the first question in the questionnaire related to gout flares, which have been identified by patients as an essential factor to consider

when measuring outcomes in chronic gout (Schumacher et al 2009). The questionnaire ended with demographic questions, which if placed at the beginning of the questionnaire may deter respondents from continuing to complete the questionnaire (Boynton & Greenhalgh, 2004). Within this questionnaire the response scales used, and the direction in which responses were presented, changed at various different points in the questionnaire. Such changes in the direction of questionnaire items and responses have been advocated as strategies to limit a response set from respondent and acquiescence bias (Bowling et al 2014).

The use of self-administered methods, such as this postal survey, have been proposed to avoid overstated estimates of HRQOL and health status that may occur in more direct modes of enquiry such as face to face or telephone interviews (Bowling et al 1999; Lyons et al 1999).

3.8 Conclusion

In conclusion, this chapter describes the prospective cohort study in which this thesis is nested. This chapter describes the study design, participants & setting, medical record review, ethical consideration, data management, questionnaire content, and attempts made to minimize bias. The aim of the following chapter is to describe the cohort of participants in this study and to investigate attrition bias and missing data.

4 Chapter Four Description of cohort and assessment of bias

4.1 Overview of chapter and aim

The previous chapter described the prospective postal questionnaire cohort study in which this thesis is nested. The aim of this chapter is to describe the cohort of participants in this study and to investigate attrition bias and missing data. Thus, within this chapter the flow of participants through the cohort will be described, and the potential for attrition bias and extent of item non-response missing data investigated.

4.2 Background

Attrition from a survey can occur when participants who are originally recruited to the survey do not then respond at every potential time-point (Lacey, Jordan & Croft, 2013; Matthews et al 2004; Twiske & de vente, 2002; Twisk, 2003). Participants may drop out of a survey by not responding to the later time-points in a survey (Twiske & de vente, 2002; Twisk, 2003) or attrition may be caused by participants responding intermittently to a survey (Cumming & Goldstein, 2016; Twiske & de vente, 2002; Twisk, 2003). Attrition in a survey has the potential to lead to attrition bias if the responders to the survey are systematically different in relation to baseline characteristics, confounders, and outcomes of interest, in comparison with those who did not respond or those who refused to take part (Chatfield, Brayne & Matthews, 2005; Kristman, Manno & Côté, 2005).

Item non-response is a form of missing data where information is not provided for certain items within a questionnaire (Fayers & Machin, 2016; Yan & Curtin, 2010). Missing data can also arise due to unit non-response; where a participant has not responded therefore a whole questionnaire has not been returned (Fayers & Machin, 2016; Yan & Curtin, 2010). Missing data can be classified as; missing completely at random (MCAR) when the absence of the data is unrelated to observed and unobserved data, missing at random (MAR) when the absence

of the data is related to observed data but not the unobserved data, or missing not at random (MNAR) when the absence of the data is related to the unobserved data (Twisk et al 2013). Attrition and missing data may have several detrimental consequences including a reduction in sample size, loss of statistical power, and biased inferences (Biering, Hjollund & Frydenberg, 2015; Deeg, 2002).

4.3 Method and Analysis Plan

4.3.1 Aim and Objectives

Aim

The aim of this chapter was to describe the flow of participants through this cohort study and investigate attrition bias and missing data.

Objectives

The objectives of this chapter were to:

1. Describe the response and the flow of participants through this cohort study.

2. Compare characteristics of responders and non-participants at the five questionnaire mail-out time points (i.e. Baseline, 6, 12, 24, 36 months).

3. Compare characteristics of participants who responded on a different number of occasions (i.e. only baseline, baseline plus one other time-point, baseline plus two other time-points, baseline plus three other time-points, all time-points).

4. Determine the number and percentage of missing item responses of key variables within the questionnaire and medical record review data; indicating the type and volume of missing item response data.

4.3.2 Data source

Data in this chapter were derived from the cohort study (Chandratre et al 2012) described in chapter three. The questionnaire was mailed to participants at baseline (October and November 2012), 6 months, 12 months, 24 months and 36 months. The data sources used in this chapter include the cohort study's mailing database, questionnaire responder database, and medical record review database.

4.3.3 Analysis plan

Flow through the cohort study was described by reporting in a flowchart the number of participants excluded before mailing, eligible for mailing, mailed, excluded during mailing and who refused or did not respond, at each time-point. Response to the questionnaire mail-outs at each time-point was calculated as the number of returned questionnaires as a percentage of the number of questionnaires mailed out to eligible participants at that time-point, thus providing an adjusted response (Bruce, Pope & Stanistreet, 2008; Sahlqvist et al 2011).

The data used to investigate attrition bias were taken from questionnaire responses at baseline and medical record review from the two years prior to baseline and included gout-specific factors, medications, comorbidities, socio-demographic factors, BMI, alcohol consumption and HRQOL scores. Characteristics at baseline are commonly used for comparison when investigating factors associated with attrition (Dumville, Torgerson & Hewitt, 2006; Fewtrell et al 2008; Lacey, Jordan & Croft, 2013; Matthews et al 2004). As attrition bias can arise when the distribution of potential confounders and outcomes are associated with whether participants remain in a study (Kristman, Manno & Côté, 2005), the variables included in this analysis of attrition bias were outcomes and potential covariates relevant to this thesis. The baseline characteristics of responders were compared with those of non-participants at each follow-up time-point. The baseline characteristics of participants

who responded only at baseline, baseline plus one other time-point, baseline plus two other time-points, baseline plus three other time-points, and all time-points were also compared.

Nominal and ordinal categorical baseline variables were analysed by undertaking frequency counts and calculating percentages. Continuous variables were analysed by calculating the mean and standard deviation.

To assess item non-response missing data, the number and percentage of responders who did not respond to each item of the GIS subscales, SF-36 PF10 and HAQ-DI was calculated at each time-point separately. The total number and percentage of responders who responded to an insufficient number of items to permit calculation of a GIS subscales score, SF-36 PF10 score and HAQ-DI score was also calculated. The frequency and percentage of missing data for gout-specific, comorbid, socio-demographic, BMI and alcohol missing data was calculated.

The analysis within this chapter was undertaken using SPSS version 24.

4.4 Results

4.4.1 Cohort response flow-chart

Figure 4.1 to Figure 4.5 display the number of participants at each stage of the study, the proportion excluded before mailing, eligible for mailing, mailed, excluded during mailing, refused or not responding, and reasons for exclusion or refusal at each time-point.

1184 participants responded at baseline (adjusted response 65.9%), with 818 (79.7%), 721 (73%), 696 (75.4%), and 605 (68.4%) participants responding at 6, 12, 24 and 36 months respectively. 1079 (91.1%) of participants responding at baseline consented to medical record review. Reasons for refusal at baseline included reporting 'no gout' (n=43), health reasons (n=9), participant not wanting to take part (n=43) and a late response (n=1). Reasons given for refusal after baseline related to health (n=12) or participants not wanting to take

part (n=126). Reasons for exclusion during mailing included death (n=15), address/contact problems (n=16) and health reasons (n=2).

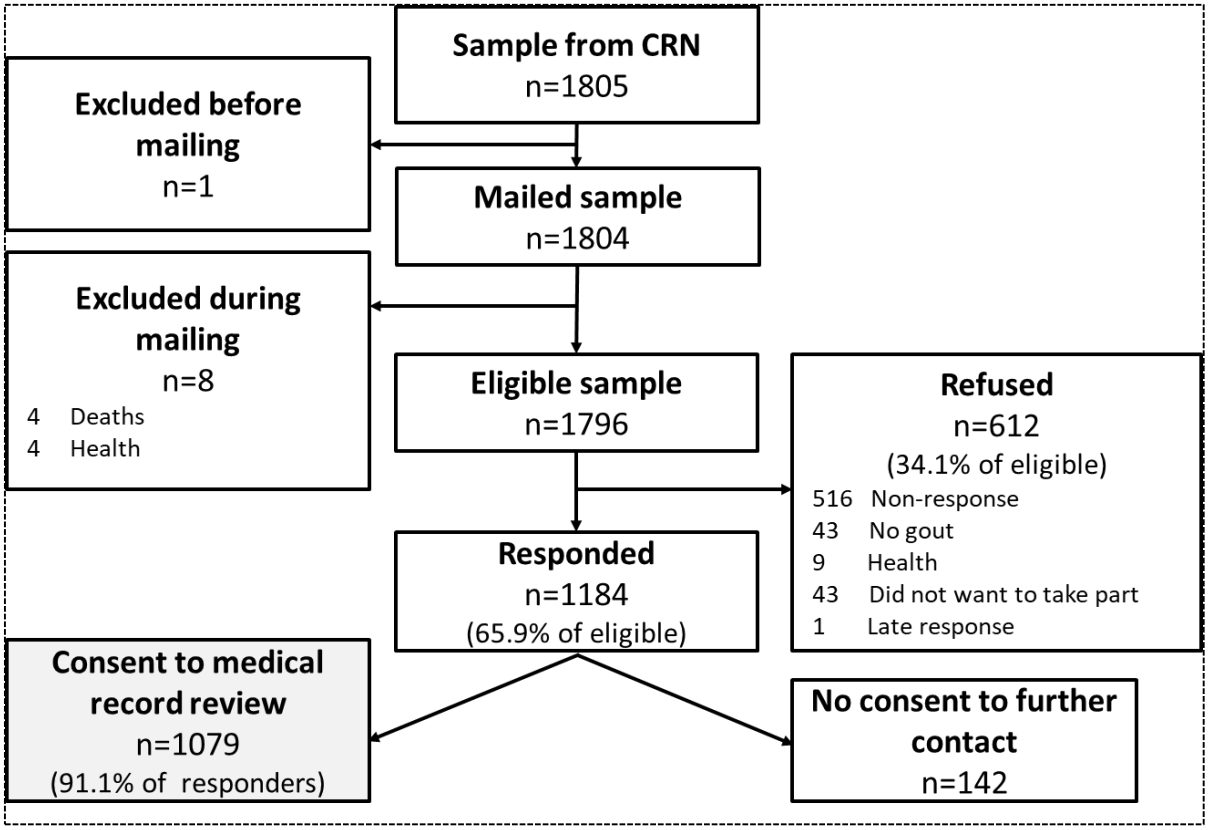


Figure 4.1 Flow diagram of exclusions, refusal, and response at baseline

CRN= Clinical Research Network

The response at baseline and number of participants consenting to medical record review have been published previously (Chandratre et al 2018).

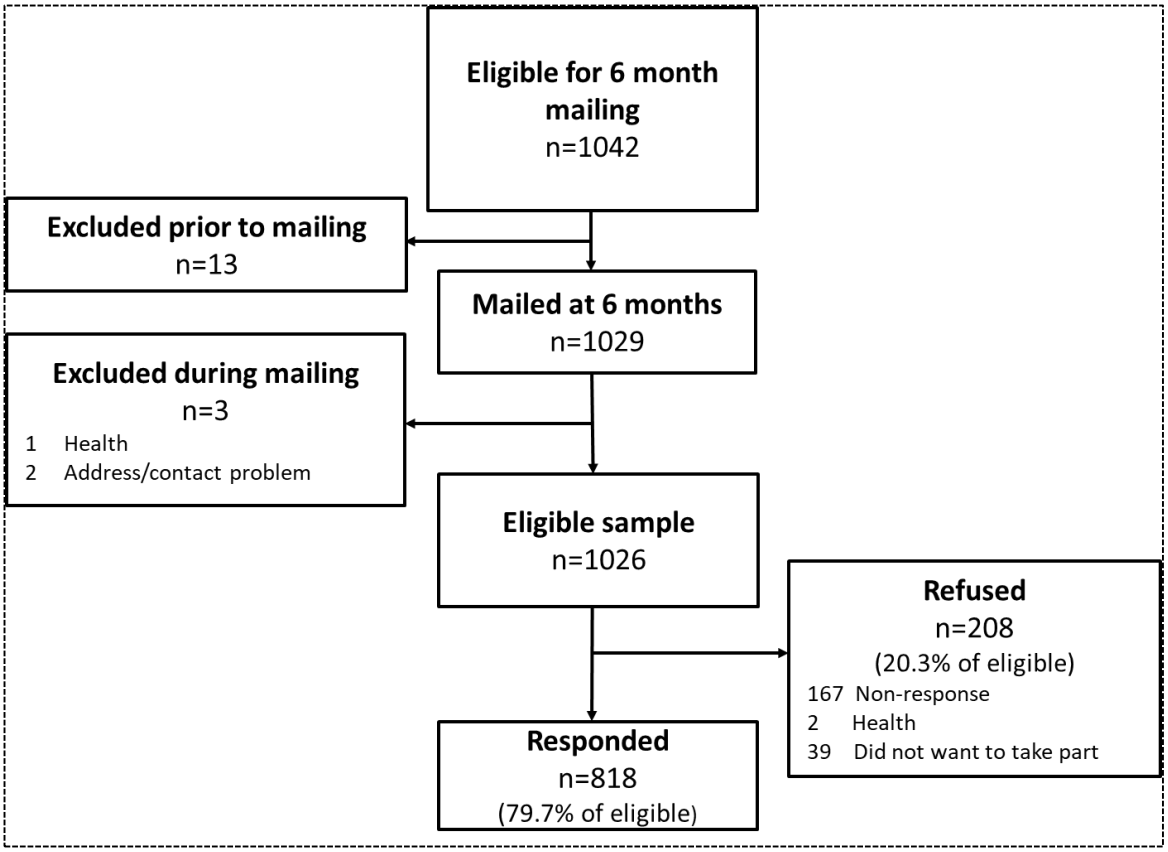


Figure 4.2 Flow diagram of exclusions, refusal and response at 6 months

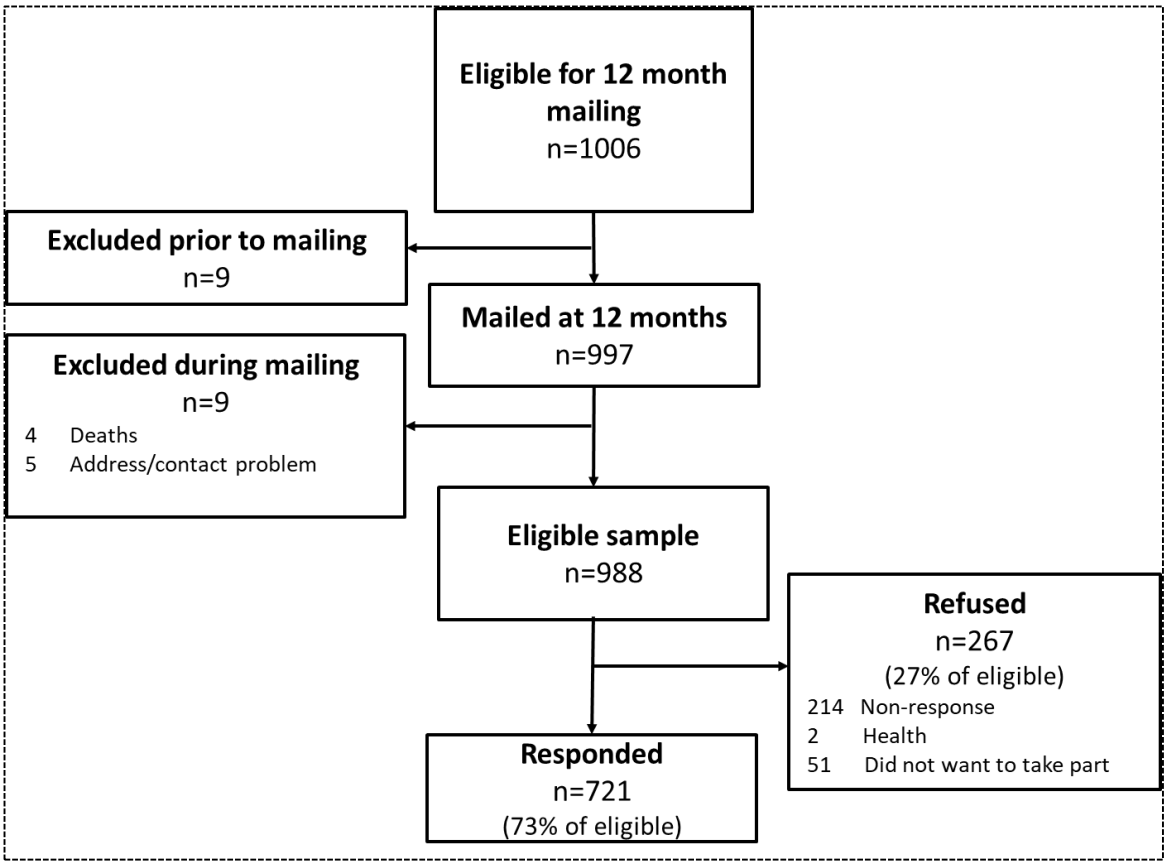


Figure 4.3 Flow diagram of exclusions, refusal and response at 12 months

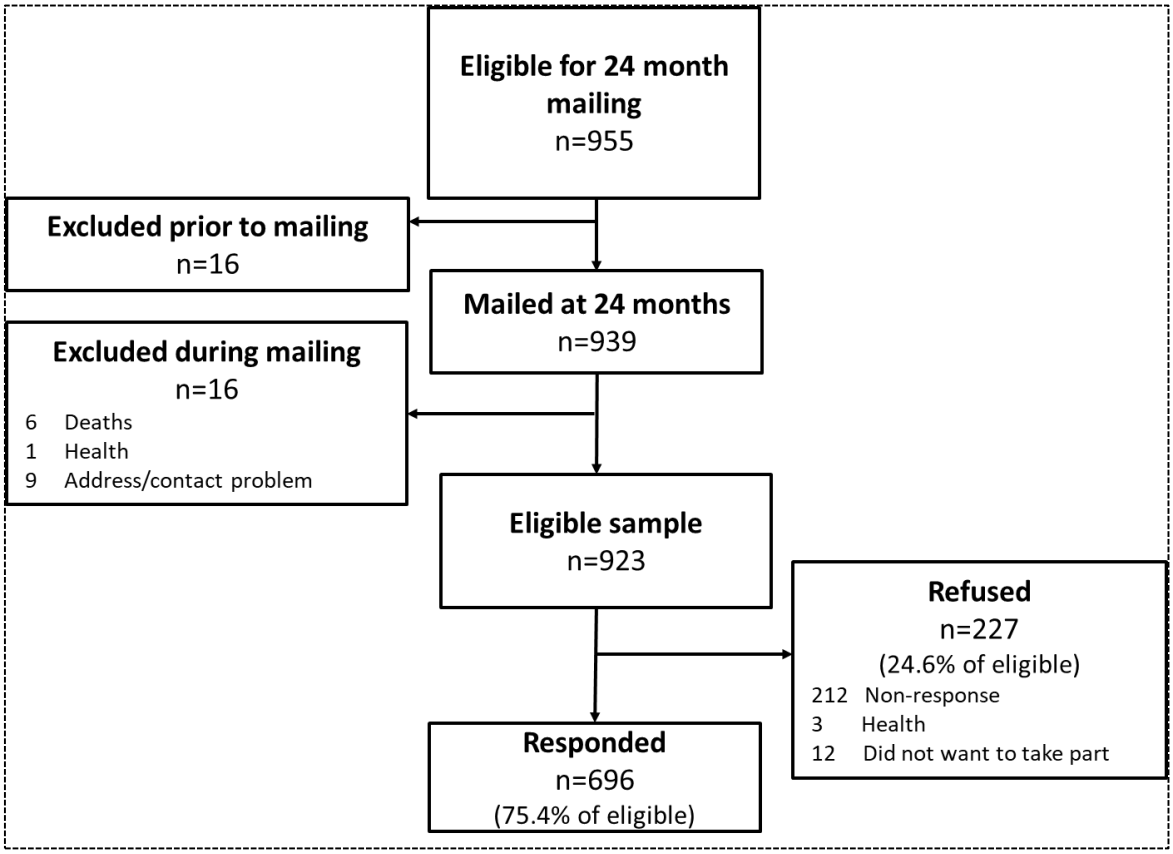


Figure 4.4 Flow diagram of exclusions, refusal and response at 24 months

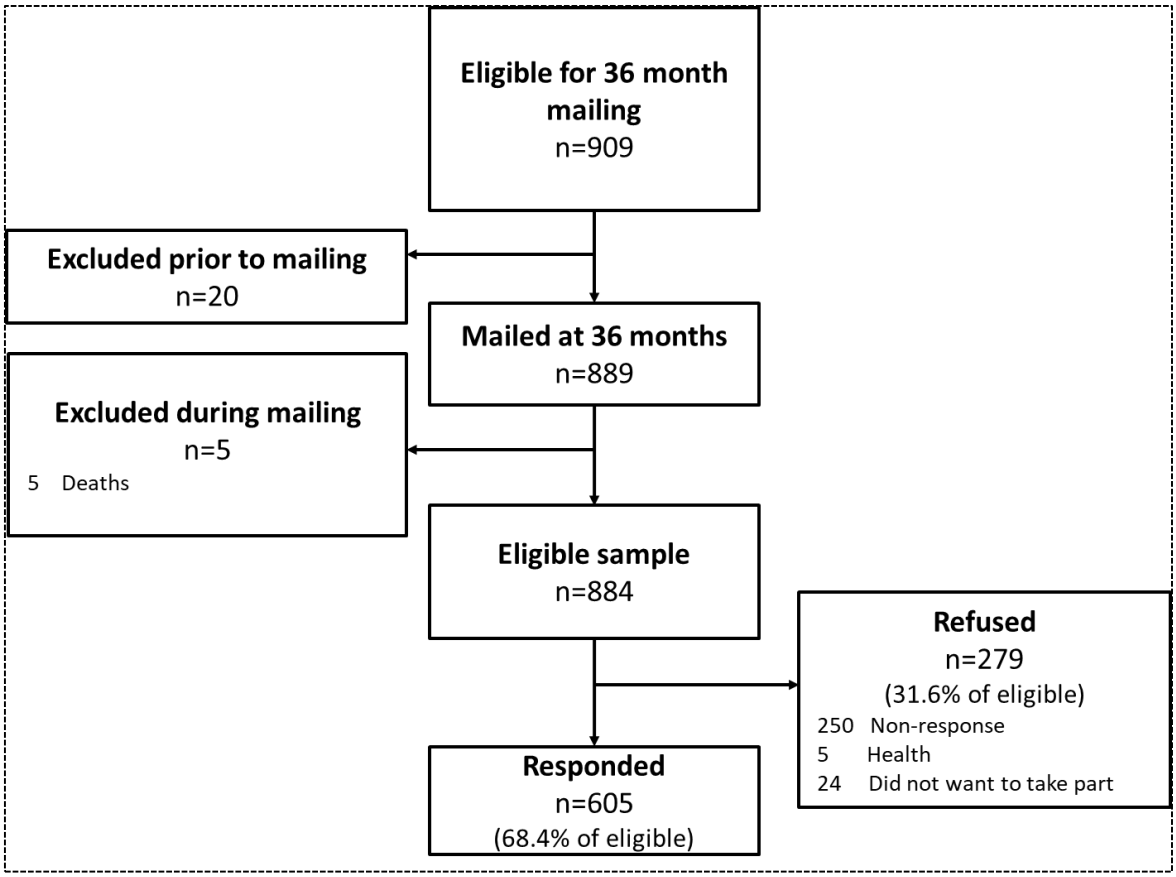


Figure 4.5 Flow diagram of exclusions, refusal and response at 36 months

4.4.2 Assessment of attrition bias

Comparison of the characteristics of responders with non-participants

Table 4.1 to Table 4.6 display the baseline characteristics of responders and non-participants at each follow-up time-point. The characteristics of the full cohort at baseline are included in each table for comparison.

Gout-specific characteristics

The proportion of individuals who had experienced two or more gout flares was slightly higher in the non-participants (ranging from 43.2 to 45.4%) in comparison with responders at six months onwards (40.0 to 40.8%) (Table 4.1). The proportion of individuals experiencing a current flare at baseline was also slightly higher in the non-participants (ranging from 12.0 to 13.6%) in comparison with responders (8.8 to 10.8%). Non-participants also had a slightly shorter mean disease duration (ranging from 10.96 to 11.28 years) than responders (ranging from 12.31 to 12.79 years). Non-participants at each time-point also had a lower proportion of individuals with a serum urate greater than 360 µmol/L (24.6 to 28.0%) when compared with responders (31.4% to 32.7%). However, the proportion of individuals with a serum urate greater than 360 µmol/L, as a proportion of those who had a serum urate record in their notes, was similar between non-participants and responders at 6, 12 and 36 months. The proportion of individuals missing a serum urate record in their medical records in the two years prior to baseline was higher in the non-participants (63.4 to 67.5%), compared with responders (57.8 to 58.8%).

Other characteristics presented in Table 4.1 did not appear to differ substantially between responders and non-participants.

Table 4.1 Gout-specific characteristics at baseline of responders and non-participants at baseline, 6, 12, 24 and 36 months

Baseline characteristic	**Baseline** characteristics	**6 Month** mail-out characteristics at baseline		**12 Month** mail-out characteristics at baseline		**24 Month** mail-out characteristics at baseline		**36 Month** mail-out characteristics at baseline	
	All responders at baseline (n=1184)	Responders (n=818)	Non-participants (n=366)	Responders (n=721)	Non-participants (n=463)	Responders (n=696)	Non-participants (n=488)	Responders (n=605)	Non-participants (n=579)
0 gout flares†	398(33.6)	284(34.7)	114(31.1)	251(34.8)	147(31.7)	248(35.6)	150(30.7)	222(36.2)	176(30.4)
1 gout flare†	231(19.5)	168(20.5)	63(17.2)	145(20.1)	86(18.6)	137(19.7)	94(19.3)	115(19.0)	116(20.0)
2 gout flares†	187 (15.8)	134(16.4)	53(14.5)	116(16.1)	71(15.3)	111(15.9)	76(15.6)	99(16.4)	88(15.2)
3 gout flares†	103(8.7)	69(8.4)	34(9.3)	64(8.9)	39(8.4)	58(8.3)	45(9.2)	57(9.4)	46(7.9)
4 gout flares†	67(5.7)	39(4.8)	28(7.7)	36(5.0)	31(6.7)	38(5.5)	29(5.9)	29(4.8)	38(6.6)
5 or more gout flares†	137(11.6)	86(10.5)	51(13.9)	78(10.8)	59(12.7)	72(10.3)	65(13.3)	57(9.4)	80(13.8)
≥2 gout flares†	494(41.7)	328(40.1)	166(45.4)	294(40.8)	200(43.2)	279(40.1)	215(44.1)	242(40.0)	252(43.5)
Disease duration years mean (SD)	11.91 (12.13)	12.32 (11.77)	10.98 (12.89)	12.31 (11.75)	11.28 (12.69)	12.36 (11.52)	11.26 (12.95)	12.79 (11.74)	10.96 (12.47)
Occurrence of current flare	132(11.1)	88(10.8)	44(12.0)	72(10.0)	60(13.0)	67(9.6)	65(13.3)	53(8.8)	79(13.6)
History of oligo/ polyarticular flares	436(36.8)	305(37.3)	131(35.8)	271(37.6)	165(35.6)	260(37.4)	176(36.1)	224(37.0)	212(36.6)
Record of tophi ◇	25(2.1)	18(2.2)	7(1.9)	17(2.4)	8(1.7)	13(1.9)	12(2.5)	12(2.0)	13(2.2)
Maximum serum urate level mean (SD) µmol/L◇*	441.36 (115.51)	438.63 (110.80)	449.23 (128.25)	439.20 (111.38)	445.55 (123.38)	437.41 (110.25)	448.06 (123.96)	437.45 (111.27)	445.96 (120.40)
Serum urate level >360 µmol/L ◇*	352(29.7)	262(32.0)	90(24.6)	236(32.7)	116(25.1)	221(31.8)	131(26.8)	190(31.4)	162(28.0)
Serum urate level >360 µmol/L as % of those with serum urate record◇*	352(76.4)	262(76.6)	90(75.6)	236(77.6)	116(73.9)	221(76.2)	131(76.6)	190(76.3)	162(76.4)
Missing serum urate record◇	723(61.1)	476(58.2)	247(67.5)	417(57.8)	306(66.1)	406(58.3)	317(65.0)	356 (58.8)	367(63.4)

Values are n(%) unless stated otherwise † **in previous 12 months** at baseline; ◇**In medical record** in the two years prior to baseline; ***highest serum urate** recorded
Baseline data for cohort have been published previously (Chandratre et al 2018).

Medications

The non-participant group at each time-point had lower proportions of individuals reporting taking allopurinol (45.9 to 48.7%) and 300mg of allopurinol (20.2 to 21.2%) at baseline compared with responders (allopurinol 56.5 to 57.5%; 300 mg 27.6% to 29.6%) (Table 4.2). This observation remained when the number of people reporting to take allopurinol, as a proportion of those who responded to the allopurinol use item, was compared between non-participants and responders. In addition, a lower proportion of individuals in the non-participant group had a prescription of allopurinol in their medical record in the two years prior to baseline (41.0 to 47.5%) in contrast with all responders at baseline (54.6%) and responders from six months onwards (60.2 to 61.3 %). A greater proportion of the non-participants either did not report an allopurinol dose or reported not knowing the allopurinol dose (52.3 to 54.6%), compared with responders (43.9 to 44.5%).

Fewer non-participants had a prescription for NSAIDs in the medical record in the two years prior to baseline (43.7 to 48.9%) in contrast with responders (57.8 to 58.5%).

Other characteristics presented in Table 4.2 did not appear to differ substantially between responders and non-participants.

Table 4.2 Medication of responders and non-participants at baseline, 6, 12, 24 and 36 months

Baseline characteristic	**Baseline** characteristics	**6 Month** mail-out characteristics at baseline		**12 Month** mail-out characteristics at baseline		**24 Month** mail-out characteristics at baseline		**36 Month** mail-out characteristics at baseline	
	All responders at baseline (n=1184)	Responders (n=818)	Non-participants (n=366)	Responders (n=721)	Non-participants (n=463)	Responders (n=696)	Non-participants (n=488)	Responders (n=605)	Non-participants (n=579)
Using allopurinol*	630(53.2)	462(56.5)	168(45.9)	411(57.0)	219(47.3)	399(57.3)	231(47.3)	348(57.5)	282(48.7)
Using allopurinol*†	630(56.3)	462(59.5)	168(49.0)	411(60.2)	219(50.1)	399(60.5)	23(50.2)	348(60.8)	282(51.5)
Allopurinol dose*									
≦100 mg	228(19.2)	162(19.8)	66(18.0)	146(20.2)	82(17.7)	134(19.2)	94(19.3)	114(18.8)	114(19.7)
150-200 mg	61(5.2)	43(5.3)	18(4.9)	33(4.6)	28(6.1)	36(5.2)	25(5.1)	32(5.3)	29(5.0)
300mg	302(25.5)	226(27.6)	76(20.8)	207(28.7)	95(20.5)	203(29.2)	99(20.2)	179(29.6)	123(21.2)
≥400 mg	29(2.5)	23(2.8)	6(1.6)	18(2.5)	11(2.4)	16(2.3)	13(2.7)	19(3.1)	10(1.7)
Don't know dose or missing	564(47.6)	364(44.5)	200(54.6)	317(44.0)	247(53.3)	307(44.1)	257(52.7)	261(43.1)	303(52.3)
Prescription Allopurinol◇	646(54.6)	496(60.6)	150(41.0)	440(61.0)	206(44.5)	419(60.2)	227(46.5)	371(61.3)	275(47.5)
Prescription Colchicine◇	345(29.1)	246(30.1)	99(27.0)	222(30.8)	123(26.6)	203(29.2)	142(29.1)	176(29.1)	169(29.2)
Prescription NSAIDs◇	634(53.5)	474(57.9)	160(43.7)	417(57.8)	217(46.9)	407(58.5)	227(46.5)	351(58.0)	283(48.9)
Prescription Diuretic◇	286(24.2)	207(25.3)	79(21.6)	183(25.4)	103(22.2)	169(24.3)	117(24.0)	136(22.5)	150(25.9)

Values are n(%) unless stated otherwise

*** Self-reported** allopurinol use and dose is taken from questionnaire responses; **†** as a proportion of participants who answered item about self-reported allopurinol; **◇ In medical record** in the two years prior to baseline.

Baseline data for cohort have been published previously (Chandratre et al 2018).

Comorbidities

In the non-participants the proportion of individuals self-reporting a diagnosis of diabetes (18.8 to 21.6%), myocardial infarction (MI) (10.6 to 12.1%), or angina (13.4 to 14.1%) was higher, and self-reporting hyperlipidaemia was lower (40.4 to 41.7%), compared with responders (diabetes 15.2 to 16.4%; MI 8.1 to 9.7%; angina 11.2 to 12 %; hyperlipidaemia 43.7 to 45.3%) (Table 4.3).

Non-participants also had slightly higher mean scores for PHQ-9 (3.99 to 4.19), GAD-7 (3.02 to 3.30), pain NRS (2.62 to 2.70), and global health NRS (2.93 to 3.01), in comparison with responders (PHQ-9 3.28 to 3.42; GAQ-7 2.55 to 2.61; pain NRS 2.03 to 2.15; global health NRS 2.38 to 2.58) (Table 4.4).

Other characteristics presented in Table 4.3 and Table 4.4 did not appear to differ substantially between responders and non-participants.

Table 4.3 Self-reported comorbidities, total comorbidities and eGFR <60 mL/min/1.73m^2 of responders and non-participants at baseline, 6, 12, 24 and 36 months

Baseline characteristic	**Baseline** characteristics	**6 Month** mail-out characteristics at baseline		**12 Month** mail-out characteristics at baseline		**24 Month** mail-out characteristics at baseline		**36 Month** mail-out characteristics at baseline	
	All responders at baseline (n=1184)	Responders (n=818)	Non-participants (n=366)	Responders (n=721)	Non-participants (n=463)	Responders (n=696)	Non-participants (n=488)	Responders (n=605)	Non-participants (n=579)
Comorbidity*									
Diabetes	205(17.3)	126(15.4)	79(21.6)	118(16.4)	87(18.8)	106(15.2)	99(20.3)	94(15.5)	111(19.2)
Cerebrovascular accident (CVA)	37(3.1)	24(2.9)	13(3.6)	24(3.3)	13(2.8)	19(2.7)	18(3.7)	14(2.3)	23(4.0)
Hypertension (HT)	731(61.7)	501(61.2)	230(62.8)	440(61.0)	291(62.9)	421(60.5)	310(63.5)	364(60.2)	367(63.4)
Transient ischaemic attack (TIA)	62(5.2)	38(4.6)	24(6.6)	33(4.6)	29(6.3)	36(5.2)	26(5.3)	30(5.0)	32(5.5)
Hyperlipidaemia (HL)	508(42.9)	360(44.0)	148(40.4)	315(43.7)	193(41.7)	311(44.7)	197(40.4)	274(45.3)	234(40.4)
Myocardial Infarction (MI)	119(10.1)	78(9.5)	41(11.2)	70(9.7)	49(10.6)	65(9.3)	54(11.1)	49(8.1)	70(12.1)
Renal failure (RF)	56(4.7)	42(5.1)	14(3.8)	36(5.0)	20(4.3)	33(4.7)	23(4.7)	21(3.5)	35(6.0)
Renal calculi	81(6.8)	53(6.5)	28(7.7)	48(6.7)	33(7.1)	45(6.5)	36(7.4)	45(7.4)	36(6.2)
Angina	147(12.4)	98(12.0)	49(13.4)	85(11.8)	62(13.4)	78(11.2)	69(14.1)	68(11.2)	79(13.6)
Total comorbidities mean (SD)†	1.6 (1.4)	1.6(1.3)	1.7(1.5)	1.6(1.4)	1.7(1.4)	1.6 (1.4)	1.7(1.4)	1.6(1.4)	1.7(1.4)
eGFR <60 mL/min/1.73m^2◇	318(26.9)	229(28.0)	89(24.3)	198(27.5)	120(25.9)	187(26.9)	131(26.8)	148(24.5)	170(29.4)

Values are n(%) unless stated otherwise ***comorbidities** self-reported in baseline questionnaire

†**total number of comorbidities** self-reported in baseline questionnaire (diabetes, hypertension, hyperlipidaemia, myocardial infarction, angina, cerebrovascular accident, transient ischaemic attack, renal failure, renal calculi)

eGFR <60 mL/min/1.73m^2 indicative of chronic kidney disease ◇ **In medical record** in the two years prior to baseline

Baseline data for cohort have been published previously (Chandratre et al 2018).

Table 4.4 Depression scores, anxiety scores, pain and global health of responders and non-participants at baseline, 6, 12, 24 and 36 months

Baseline characteristic	**Baseline** characteristics	**6 Month** mail-out characteristics at baseline		**12 Month** mail-out characteristics at baseline		**24 Month** mail-out characteristics at baseline		**36 Month** mail-out characteristics at baseline	
	All responders at baseline (n=1184)	Responders (n=818)	Non-participants (n=366)	Responders (n=721)	Non-participants (n=463)	Responders (n=696)	Non-participants (n=488)	Responders (n=605)	Non-participants (n=579)
Depression (PHQ-9 category)									
Minimal	763(64.4)	547(66.9)	216(59.0)	490(68.0)	273(59.0)	475(68.2)	288(59.0)	430(71.1)	333(57.5)
Mild	148(12.5)	108(13.2)	40(10.9)	83(11.5)	65(14.0)	85(12.2)	63(12.9)	71(11.7)	77(13.3)
Moderate	65(5.5)	44(5.4)	21(5.7)	42(5.8)	23(5.0)	38(5.5)	27(5.5)	30(5.0)	35(6.0)
Moderately severe	40(3.4)	22(2.7)	18(4.9)	19(2.6)	21(4.5)	17(2.4)	23(4.7)	15(2.5)	25(4.3)
Severe	26(2.2)	15(1.8)	11(3.0)	16(2.2)	10(2.2)	14(2.0)	12(2.5)	14(2.3)	12(2.1)
PHQ-9 score mean(SD)	3.64(5.22)	3.41(4.94)	4.19(5.83)	3.42(5.08)	3.99(5.44)	3.35(4.97)	4.07(5.57)	3.28(4.97)	4.06(5.48)
Anxiety (GAD-7 category)									
Minimal	844(71.3)	607(74.2)	237(64.8)	533(73.9)	311(67.2)	518(74.4)	326(66.8)	454(75.0)	390(67.4)
Mild	141(11.9)	99(12.1)	42(11.5)	90(12.5)	51(11.0)	91(13.1)	50(10.2)	75(12.4)	66(11.4)
Moderate	64(5.4)	39(4.8)	25(6.8)	36(5.0)	28(6.0)	30(4.3)	34(7.0)	25(4.1)	39(6.7)
Severe	45(3.8)	28(3.4)	17(4.6)	24(3.3)	21(4.5)	22(3.2)	23(4.7)	22(3.6)	23(4.0)
GAD-7 score mean(SD)	2.79(4.49)	2.58(4.21)	3.30(5.07)	2.61(4.26)	3.08(4.84)	2.55(4.14)	3.15(4.96)	2.58(4.25)	3.02(4.74)
Body pain	651(54.9)	465(56.8)	186(50.8)	415(57.6)	236(51.0)	383(55.0)	268(54.9)	339(56.0)	312(53.9)
NRS pain in last week mean (SD)	2.32(2.85)	2.15(2.72)	2.70(3.11)	2.11(2.74)	2.65(2.99)	2.08(2.73)	2.67(2.99)	2.03(2.68)	2.62(3.00)
NRS global health mean (SD)	2.68(2.77)	2.58(2.70)	2.93(2.92)	2.51(2.70)	2.97(2.87)	2.48(2.68)	2.98(2.88)	2.38(2.68)	3.01(2.84)

Values are n(%) unless stated otherwise

PHQ-9 score ranges from 0 to 27 Minimal depression 0-4, Mild 5-9, Moderate 10-14, Moderately Severe 15-19, Severe 20-27; **GAD-7 score** ranges from 0 to 21 Minimal anxiety 0-4, Mild 5-9, Moderate 10-14, Severe 15-21; **Body pain** (including ache or discomfort or stiffness) for one day or longer in the 4 weeks prior to baseline**; NRS pain in last week** ranges from 0 (no pain) to 10 (pain as bad as it can be)**; NRS global health** ranges from 0 (very well) to 10 (very poor health).

Baseline data for cohort have been published previously (Chandratre et al 2018).

HRQOL measures; GIS subscales, SF-36 PF10 and HAQ-DI

The mean GIS unmet treatment need (UTN) score of non-participants was higher (ranging from 36.64 to 37.89) thus indicating slightly worse gout-specific HRQOL, in contrast with the mean GIS UTN scores of responders (30.57 to 31.60) (Table 4.5). However, there did not appear to be substantial differences in the other GIS subscales scores between responders and non-participants.

The mean SF-36 PF10 scores at each time-point were lower for non-participants (69.54 to 70.31) thus indicating worse HRQOL, in comparison with responders (78.79 to 81.52).

At each time-point, the mean HAQ-DI scores were higher for non-participants (0.63 to 0.64) thus indicating greater activity limitation in comparison with responders (0.39 to 0.46).

Table 4.5 GIS subscales, SF-36 PF10 and HAQ-DI scores of responders and non-participants at baseline, 6, 12, 24 and 36 months

Baseline characteristic	**Baseline** characteristics	**6 Month** mail-out characteristics at baseline		**12 Month** mail-out characteristics at baseline		**24 Month** mail-out characteristics at baseline		**36 Month** mail-out characteristics at baseline	
	All responders at baseline (n=1184)	Responders (n=818)	Non-participants (n=366)	Responders (n=721)	Non-participants (n=463)	Responders (n=696)	Non-participants (n=488)	Responders (n=605)	Non-participants (n=579)
GIS Concern overall (CO)	48.65(28.33)	47.68(28.10)	50.89(28.77)	47.65(27.93)	50.25(28.92)	47.96(28.40)	49.67(28.23)	47.52(28.14)	49.87(28.51)
GIS Medication side effects (MSE)	40.45(26.33)	40.20(25.75)	41.05(27.67)	40.01(25.68)	41.17(27.36)	40.30(26.40)	40.68(26.24)	40.59(26.46)	40.31(26.20)
GIS Unmet treatment need (UTN)	33.46(20.57)	31.60(20.11)	37.89(20.99)	31.35(20.09)	36.93(20.90)	30.72(19.65)	37.61(21.24)	30.57(19.26)	36.64(21.49)
GIS Wellbeing during an attack (WBDA)	45.19(26.41)	45.32(25.94)	44.88(27.49)	45.24(25.93)	45.12(27.19)	45.44(26.36)	44.83(26.51)	45.94(26.45)	44.38(26.37)
GIS Concern during an attack (CDA)	40.13(24.35)	39.59(23.68)	41.38(25.83)	39.40(23.83)	41.32(25.15)	39.55(24.40)	40.99(24.27)	39.40(23.93)	40.92(24.79)
SF-36 PF10	75.91(26.12)	78.79(24.12)	69.57(29.09)	79.45(24.00)	70.31(28.30)	80.25(23.89)	69.54(27.92)	81.53(22.78)	69.78(28.10)
HAQ-DI	0.51(0.71)	0.46(0.67)	0.63(0.80)	0.43(0.64)	0.63(0.80)	0.41(0.63)	0.64(0.80)	0.39(0.62)	0.64(0.78)

Values are mean (SD)

Each **GIS subscale** scored from **0 to 100;** higher scores on each scale indicating a greater impact of gout on HRQOL/ worse HRQOL.

SF-36 PF10 scored from **0 to 100**; higher score indicating performs all types of physical activities including the most vigorous without limitations due to health.

HAQ-DI scored from **0 to 3**; higher score indicating greater activity limitation.

Baseline data for cohort have been published previously (Chandratre et al 2018).

Socio-demographic characteristics

As displayed in Table 4.6 responders were more likely to be male (86.2 to 89.4% male), attended further education (23.6 to 26.0%) and drink alcohol either daily or almost daily (24.4 to 26.9%), compared with non-participants (male 77.5 to 78.1%; further education 15.3 to 16.2%; alcohol daily 18.4 to 19.9%). Fewer responders were over the age of 80 (8.6 to 10.5%) and classified as most deprived (27.1 to 29.0%), in comparison with non-participants (over 80 16.8 to 18.2%; most deprived 35.4 to 36.5%).

BMI and alcohol frequency

As displayed in Table 4.7 responders were more likely to be drink alcohol either daily or almost daily (24.4 to 26.9%), compared with non-participants (alcohol daily 18.4 to 19.9%). No marked differences were observed in the BMI of responders and non-participants.

Table 4.6 Socio-demographic characteristics of responders and non-participants at baseline, 6, 12, 24 and 36 months

Baseline characteristic	**Baseline** characteristics	**6 Month** mail-out characteristics at baseline		**12 Month** mail-out characteristics at baseline		**24 Month** mail-out characteristics at baseline		**36 Month** mail-out characteristics at baseline	
	All responders at baseline (n=1184)	Responders (n=818)	Non-participants (n=366)	Responders (n=721)	Non-participants (n=463)	Responders (n=696)	Non-participants (n=488)	Responders (n=605)	Non-participants (n=579)
Male	990(83.6)	705(86.2)	285(77.9)	629(87.2)	361(78.0)	609(87.5)	381(78.1)	541(89.4)	449(77.5)
Age mean (SD)	65.61(12.49)	65.43(11.96)	66.01(13.58)	65.35(11.68)	66.01(13.66)	65.09(11.48)	66.35(13.78)	64.58(11.48)	66.68(13.40)
Age categories									
<40	26(2.2)	17(2.1)	9(2.5)	16(2.2)	10(2.2)	15(2.2)	11(2.3)	15(2.5)	11(1.9)
40-49.9	116(9.8)	73(8.9)	43(11.7)	59(8.2)	57(12.3)	56(8.0)	60(12.3)	50(8.3)	66(11.4)
50-59.9	210(17.7)	151(18.5)	59(16.1)	137(19.0)	73(15.8)	132(19.0)	78(16.0)	121(20.0)	89(15.4)
60-69.9	343(29.0)	253(30.9)	90(24.6)	227(31.5)	116(25.1)	232(33.3)	111(22.7)	206(34.0)	137(23.7)
70-79.9	339(28.6)	238(29.1)	101(27.6)	210(29.1)	129(27.9)	200(28.7)	139(28.5)	161(26.6)	178(30.7)
>80	150(12.7)	86(10.5)	64(17.5)	72(10.0)	78(16.8)	61(8.8)	89(18.2)	52(8.6)	98(16.9)
Neighbour deprivation status†									
Most deprived	369(31.2)	237(29.0)	132(36.1)	200(27.7)	169(36.5)	193(27.7)	176(36.1)	164(27.1)	205(35.4)
Middle	405(34.2)	289(35.3)	116(31.7)	257(35.6)	148(32.0)	260(37.4)	145(29.7)	45(38.3)	173(29.9)
Least deprived	410(34.6)	292(35.7)	118(32.2)	264(36.6)	146(31.5)	243(34.9)	167(34.2)	37(34.5)	201(34.7)
Ethnic origin									
White UK/European	1126(95.1)	789(96.5)	337(92.1)	698(96.8)	428(92.4)	670(96.3)	456(93.4)	589(97.4)	537(92.7)
Asian	16(1.4)	9(1.1)	7(1.9)	8(1.1)	8(1.7)	10(1.4)	6(1.2)	6(1)	10(1.7)
Afro Caribbean	2(0.2)	2(0.2)	0(0)	1(0.1)	1(0.2)	1(0.1)	1(0.2)	1(0.2)	1(0.2)
African	2(0.2)	1(0.1)	1(0.3)	0(0)	2(0.4)	0(0)	2(0.4)	3(0.5)	2(0.3)
Chinese	1(0.1)	0(0)	1(0.3)	0(0)	1(0.2)	0(0)	1(0.2)	0(0)	1(0.2)
Other	7(0.6)	4(0.5)	3(0.8)	4(0.6)	3(0.6)	4(0.6)	3(0.6)	0(0)	4(0.7)

Values are n(%) unless stated otherwise

† calculated using tertiles of indices of multiple deprivation (IMD) Baseline data for cohort have been published previously (Chandratre et al 2018).

Table 4.6 cont. Socio-demographic characteristics of responders and non-participants at baseline, 6, 12, 24 and 36 months

Baseline characteristic	**Baseline** characteristics	**6 Month** mail-out characteristics at baseline		**12 Month** mail-out characteristics at baseline		**24 Month** mail-out characteristics at baseline		**36 Month** mail-out characteristics at baseline	
	All responders at baseline (n=1184)	Responders (n=818)	Non-participants (n=366)	Responders (n=721)	Non-participants (n=463)	Responders (n=696)	Non-participants (n=488)	Responders (n=605)	Non-participants (n=579)
Relationship status									
Married	809(68.4)	570(69.7)	239(65.3)	495(68.7)	314(67.8)	491(70.5)	318(65.2)	432(71.4)	377(65.1)
Widowed	114(9.6)	74(9.0)	40(10.9)	67(9.3)	47(10.2)	55(7.9)	59(12.1)	45(7.4)	69(11.9)
Cohabiting	73(6.2)	53(6.5)	20(5.5)	43(6.0)	30(6.5)	43(6.2)	30(6.1)	37(6.1)	36(6.2)
Divorced	69(5.8)	44(5.4)	25(6.8)	45(6.2)	24(5.2)	38(5.5)	31(6.4)	34(5.6)	35(6.0)
Separated	22(1.9)	13(1.6)	9(2.5)	14(1.9)	8(1.7)	12(1.7)	10(2.0)	9(1.5)	13(2.2)
Single	78(6.6)	54(6.6)	24(6.6)	48(6.7)	30(6.5)	48(6.9)	30(6.1)	42(6.9)	36(6.2)
Attendance at further education	249(21.0)	193(23.6)	56(15.3)	175(24.3)	74(16.0)	170(24.4)	79(16.2)	157(26.0)	92(15.9)

Values are n(%) unless stated otherwise

Baseline data for cohort have been published previously (Chandratre et al 2018).

Table 4.7 BMI and alcohol frequency of responders and non-participants at baseline, 6, 12, 24 and 36 months

Baseline characteristic	**Baseline** characteristics	**6 Month** mail-out characteristics at baseline		**12 Month** mail-out characteristics at baseline		**24 Month** mail-out characteristics at baseline		**36 Month** mail-out characteristics at baseline	
	All responders at baseline (n=1184)	Responders (n=818)	Non-participants (n=366)	Responders (n=721)	Non-participants (n=463)	Responders (n=696)	Non-participants (n=488)	Responders (n=605)	Non-participants (n=579)
BMI kg/m^2 mean (SD)	29.13(5.11)	29.30(5.24)	28.73(4.78)	29.08(5.06)	29.21(5.20)	29.21(5.04)	29.01(5.21)	29.32(5.06)	28.92(5.15)
BMI categories									
<18.5 kg/m^2	1(0.1)	1(0.1)	0(0)	1(0.1)	0(0)	1(0.1)	0(0)	1(0.2)	0(0)
18.5-24.9 kg/m^2	219(18.5)	144(17.6)	75(20.5)	134(18.6)	85(18.4)	122(17.5)	97(19.9)	106(17.5)	113(19.5)
25-29.9 kg/m^2	511(43.2)	358(43.8)	153(41.8)	323(44.8)	188(40.6)	313(45.0)	198(40.6)	262(43.3)	249(43.0)
30-34.9 kg/m^2	260(22.0)	188(23.0)	72(19.7)	160(22.2)	100(21.6)	155(22.3)	105(21.5)	147(24.3)	113(19.5)
35-39.9 kg/m^2	90(7.6)	61(7.5)	29(7.9)	48(6.7)	42(9.1)	54(7.8)	36(7.4)	46(7.6)	44(7.6)
≥40 kg/m^2	37(3.1)	31(3.8)	6(1.6)	25(3.5)	12(2.6)	22(3.2)	15(3.1)	21(3.5)	16(2.8)
Alcohol frequency									
Never	113(9.5)	73(8.9)	40(10.9)	57(7.9)	56(12.1)	57(8.2)	56(11.5)	46(7.6)	67(11.6)
Special occasions	155(13.1)	98(12.0)	57(15.6)	86(11.9)	69(14.9)	81(11.6)	74(15.2)	62(10.2)	93(16.1)
1 to 3 times per month	109(9.2)	77(9.4)	32(8.7)	71(9.8)	38(8.2)	67(9.6)	42(8.6)	60(9.9)	49(8.5)
Once or twice per week	254(21.5)	180(22.0)	74(20.2)	156(21.6)	98(21.2)	148(21.3)	106(21.7)	132(21.8)	122(21.1)
3 to 4 times per week	263(22.2)	183(22.4)	80(21.9)	163(22.6)	100(21.6)	153(22.0)	110(22.5)	139(23.0)	124(21.4)
Daily or almost daily	273(23.1)	200(24.4)	73(19.9)	182(25.2)	91(19.7)	183(26.3)	90(18.4)	163(26.9)	110(19.0)

Values are n(%) unless stated otherwise

Baseline data for cohort have been published previously (Chandratre et al 2018).

Comparison of the characteristics of participants responding on a different number of occasions

Table 4.8 to Table 4.13 display the baseline characteristics of participants who responded on a different number of occasions (i.e. only baseline, baseline plus one other time-point, baseline plus two other time-points, baseline plus three other time-points, and all time points).

Gout-specific characteristics

Participants who responded only at baseline had a shorter disease duration (10.98 years) and fewer of these participants had a serum urate greater than 360 μmol/L (21.8%) compared with participants who responded more often (Table 4.8). However, there was less difference in the proportion of individuals with a serum urate greater than 360 μmol/L, as a proportion of those who had a serum urate record in their notes. The proportion of individuals missing a serum urate record in their medical records in the two years prior to baseline was higher in participants who responded only at baseline (70.7%), compared with participants who responded more often.

Participants who responded at all five time-points had the longest disease duration (12.83 years) and fewer of these participants reported experiencing two or more flares (37.4%) or a current flare whilst completing the baseline questionnaire (7.6%) compared with participants who responded less often.

Other characteristics presented in Table 4.8 did not appear to differ substantially between the groups of participants who responded on a different number of occasions.

Table 4.8 Gout-specific characteristics of participants responding on different number of occasions

Baseline characteristic	All responders **at baseline** (n=1184)	Responded **only at baseline** (n=280)	Responders at baseline and at least **one other time-point** (n= 904)	Responders at baseline and at least **two other time-points** (n=771)	Responders at baseline and at least **three other time-points** (n=655)	Responders at baseline and **all time-points** (n=510)
0 gout flares[†]	398(33.6)	91(32.5)	307(34.0)	267(34.6)	236(36.0)	195(38.2)
1 gout flare[†]	231(19.5)	51(18.2)	180(19.9)	152(19.7)	131(20.0)	102(20.0)
2 gout flares[†]	187 (15.8)	38(13.6)	149(16.5)	123(16.0)	106(16.2)	82(16.1)
3 gout flares[†]	103(8.7)	22(7.9)	81(9.0)	69(8.9)	55(8.4)	43(8.4)
4 gout flares[†]	67(5.7)	18(6.4)	49(5.4)	40(5.2)	31(4.7)	22(4.3)
5 or more gout flares[†]	137(11.6)	40(14.3)	97(10.7)	85(11.0)	67(10.2)	44(8.6)
≥2 gout flares[†]	494(41.7)	118(42.1)	376(41.6)	317(41.1)	259(39.5)	191(37.4)
Disease duration years mean (SD)	11.91 (12.13)	10.98 (13.25)	12.19 (11.77)	12.34 (11.71)	12.54 (11.55)	12.83 (11.76)
Occurrence of current flare	132(11.1)	34(12.1)	98(10.8)	81(10.5)	62(9.5)	39(7.6)
History of oligo/ polyarticular flares	436(36.8)	97(34.6)	339(37.5)	291(37.7)	248(37.9)	182(35.7)
Record of tophi [◇]	25(2.1)	5(1.8)	20(2.2)	18(2.3)	13(2.0)	9(1.8)
Maximum serum urate level mean (SD) µmol/L[◇]*	441.36 (115.51)	445.55 (132.12)	440.46 (111.77)	439.95 (111.17)	438.08 (109.63)	431.91 (110.42)
Serum urate level >360 µmol/L [◇]*	352(29.7)	61(21.8)	291(32.2)	248(32.2)	209(31.9)	161(31.6)
Serum urate level >360 µmol/L as % of those with serum urate record[◇]*	352(76.4)	61(74.4)	291(76.8)	248(77.5)	209(76.8)	161(75.2)
Missing serum urate record[◇]	723(61.1)	198(70.7)	525(58.1)	451(58.5)	383(58.5)	296(58.0)

Values are n(%) unless stated otherwise

[†] **in previous 12 months** at baseline; [◇]**In medical record** in the two years prior to baseline; ***highest serum urate** recorded

Baseline data for cohort have been published previously (Chandratre et al 2018).

Medication

Fewer participants who responded only at baseline reported allopurinol use at baseline (46.1%) with fewer taking a dose of 300 mg (17.9%) compared to participants who responded at more time-points (Table 4.9). The number of people reporting to take allopurinol, as a proportion of those who responded to the allopurinol use item, was also less in the participants who responded only at baseline (50.8%), compared with other participants. Also, fewer participants who responded only at baseline had prescriptions for allopurinol (39.6%), colchicine (24.6%), NSAIDs (39.4%), and diuretics (21.1%) in their medical records compared with other participants. Conversely the group of participants who responded at all five time-points had more participants reporting to take allopurinol (59.0%) and with a prescription for allopurinol (62.5%).

Other characteristics presented in Table 4.9 did not appear to differ greatly between the groups of participants who responded on a different number of occasions.

Table 4.9 Medications of participants responding on different number of occasions

Baseline characteristic	All responders **at baseline** (n=1184)	Responded **only at baseline** (n=280)	Responders at baseline and at least **one other time-point** (n= 904)	Responders at baseline and at least **two other time-points** (n=771)	Responders at baseline and at least **three other time-points** (n=655)	Responders at baseline and **all time-points** (n=510)
Using allopurinol *	630(53.2)	129(46.1)	501(55.4)	440(57.1)	378(57.7)	301(59.0)
Using allopurinol *†	630(56.3)	129(50.8)	501(58.4)	440(60.3)	378(61.0)	301(62.2)
Reported allopurinol dose						
≤100mg	228(19.3)	59(21.0)	169(18.6)	153(19.8)	129(19.7)	105(20.5)
150-200 mg	61(5.2)	14(5.0)	47(5.2)	39(5.1)	33(5.0)	25(4.9)
300 mg	302(25.5)	50(17.9)	252(27.9)	220(28.5)	190(29.0)	153(30.0)
≥400mg	29(2.4)	5(1.8)	24(2.7)	22(2.9)	17(2.6)	13(2.6)
Missing or don't know dose	564(47.6)	152(54.3)	412(45.6)	337(43.7)	286(43.7)	214(42.0)
Prescription Allopurinol◇	646(54.6)	111(39.6)	535(59.2)	470(61.1)	402(61.4)	319(62.5)
Prescription Colchicine◇	345(29.1)	69(24.6)	276(30.5)	236(30.6)	191(29.2)	144(28.2)
Prescription NSAIDs◇	634(53.5)	111(39.4)	523(57.9)	448(58.7)	382(58.3)	296(58.0)
Prescription Diuretic◇	286(24.2)	59(21.1)	227(25.1)	191(24.8)	158(24.1)	119(23.3)

Values are n(%) unless stated otherwise

***Self-reported allopurinol use** and dose is taken from questionnaire responses; **† as a proportion of participants who answered item** about self-reported allopurinol**;** ◇ **In medical record** in the two years prior to baseline

Baseline data for cohort have been published previously (Chandratre et al 2018).

Comorbidities

The group who responded only at baseline had a higher proportion of participants reporting diabetes (22.5%) (Table 4.10) and a higher PHQ-9 (3.90), GAD-7 (3.04), NRS pain (2.73) and NRS global health (2.91) score compared to participants who responded more often (Table 4.11). The group of participants who responded at all five time-points had the lowest mean PHQ-9 (2.96), GAD-7 (2.32), NRS pain (1.90) and NRS global health (2.31) scores.

Other characteristics presented in Table 4.10 and Table 4.11 did not appear to differ substantially between the groups of participants who responded on a different number of occasions.

Table 4.10 Self-reported comorbidities, total comorbidities and eGFR <60 mL/min/1.73m^2 of participants responding on different number of occasions

Baseline characteristic	All responders at **baseline** (n=1184)	Responded **only at baseline** (n=280)	Responders at baseline and at least **one other time-point** (n=904)	Responders at baseline and at least **two other time-points** (n=771)	Responders at baseline and at least **three other time-points** (n=655)	Responders at baseline and **all time-points** (n=510)
Comorbidity*						
Diabetes	205(17.3)	63(22.5)	142(15.7)	119(15.4)	102(15.6)	81(15.9)
Cerebrovascular accident (CVA)	37(3.1)	9(3.2)	28(3.1)	24(3.1)	18(2.7)	11(2.2)
Hypertension (HT)	731(61.7)	177(63.2)	554(61.3)	467(60.6)	397(60.6)	308(60.4)
Transient ischaemic attack (TIA)	62(5.2)	17(6.1)	45(5.0)	36(4.7)	31(4.7)	25(4.9)
Hyperlipidaemia (HL)	508(42.9)	109(38.9)	399(44.1)	340(44.1)	287(43.8)	234(45.9)
Myocardial Infarction (MI)	119(10.1)	30(10.7)	89(9.8)	71(9.2)	58(8.9)	44(8.6)
Renal failure (RF)	56(4.7)	11(3.9)	45(5.0)	37(4.8)	31(4.7)	19(3.7)
Renal calculi	81(6.8)	24(8.6)	57(6.3)	51(6.6)	46(7.0)	37(7.3)
Angina	147(12.4)	31(11.1)	116(12.8)	90(11.7)	69(10.5)	54(10.6)
Total comorbidities mean (SD)†	1.6(1.4)	1.7 (1.4)	1.6(1.4)	1.6(1.4)	1.6(1.3)	1.6(1.3)
eGFR <60 mL/min/1.73m^2◇	318(26.9)	71(25.4)	247(27.3)	206(26.7)	173(25.4)	136(26.7)

Values are n(%) unless stated otherwise ***comorbidities** self-reported in baseline questionnaire

† **total number of comorbidities** self-reported in baseline questionnaire (diabetes, hypertension, hyperlipidaemia, myocardial infarction, angina, cerebrovascular accident, transient ischaemic attack, renal failure, renal calculi)

eGFR <60 mL/min/1.73m^2 indicative of chronic kidney disease ◇ **In medical record** in the two years prior to baseline

Baseline data for cohort have been published previously (Chandratre et al 2018).

Table 4.11 Depression scores, anxiety scores, pain and global health of participants responding on different number of occasions

Baseline characteristic	All responders at **baseline** (n=1184)	Responded **only at baseline** (n=280)	Responders at baseline and at least **one other time-point** (n=904)	Responders at baseline and at least **two other time-points** (n=771)	Responders at baseline and at least **three other time-points** (n=655)	Responders at baseline and **all time-points** (n=510)
Depression (PHQ-9 category)						
Minimal	763(64.4)	169(60.4)	594(65.7)	525(68.1)	451(68.9)	372(72.9)
Mild	148(12.5)	32(11.4)	116(12.8)	92(11.9)	79(12.1)	60(11.8)
Moderate	65(5.5)	11(3.9)	54(6.0)	45(5.8)	35(5.3)	20(3.9)
Moderately severe	40(3.4)	16(5.7)	24(2.7)	20(2.6)	18(2.7)	11(2.2)
Severe	26(2.2)	6(2.1)	20(2.2)	17(2.2)	13(2.0)	9(1.8)
PHQ-9 score mean(SD)	3.64(5.22)	3.90(5.59)	3.56(5.12)	3.44(5.07)	3.35(4.99)	2.96(4.61)
Anxiety (GAD-7 category)						
Minimal	844(71.3)	182(65.0)	662(73.2)	570(73.9)	487(74.4)	393(77.1)
Mild	141(11.9)	28(10.0)	113(12.5)	98(12.7)	84(12.8)	60(11.8)
Moderate	64(5.4)	19(6.8)	45(5.0)	37(4.8)	30(4.6)	18(3.5)
Severe	45(3.8)	11(3.9)	34(3.8)	26(3.4)	21(3.2)	15(2.9)
GAD-7 score mean (SD)	2.79(4.49)	3.04(4.89)	2.72(4.37)	2.62(4.25)	2.55(4.13)	2.32(3.96)
Body pain	651(54.9)	137(48.9)	514(56.9)	438(56.8)	369(56.4)	281(55.1)
NRS pain in last week mean (SD)	2.32(2.85)	2.73(3.10)	2.19(2.76)	2.13(2.75)	2.08(2.74)	1.90(2.55)
NRS global health mean (SD)	2.68(2.77)	2.91(2.90)	2.62(2.73)	2.50(2.69)	2.47(2.69)	2.31(2.62)

Values are n(%) unless stated otherwise

PHQ-9 score ranges from 0 to 27 Minimal depression 0-4, Mild 5-9, Moderate 10-14, Moderately Severe 15-19, Severe 20-27; **GAD-7 score** ranges from 0 to 21 Minimal anxiety 0-4, Mild 5-9, Moderate 10-14, Severe 15-21; **body pain** (including ache or discomfort or stiffness) for one day or longer in the 4 weeks prior to baseline**; NRS pain in last week** ranges from 0 (no pain) to 10 (pain as bad as it can be)**; NRS global health** ranges from 0 (very well) to 10 (very poor health).

Baseline data for cohort have been published previously (Chandratre et al 2018).

HRQOL measures; GIS subscales, SF-36 PF10 and HAQ-DI

The group of participants who responded only at baseline had a slightly higher mean GIS UTN score (38.27) indicating worse gout-specific HRQOL, along with a lower SF-36 PF10 score (67.98) indicating worse HRQOL, and a higher HAQ-DI score (0.65) indicating greater activity limitation in comparison with other participants (Table 4.12).

The group which responded at all five time-points had slightly lower GIS subscale scores (30.05 to 45.43) indicating slightly better gout-specific HRQOL, compared with scores for the whole cohort at baseline (GIS subscales ranging from 33.46 to 48.65), with the GIS CO and UTN displaying more difference compared with other GIS subscale scores. The group which responded at all five time-points also had a higher score for SF-36 PF10 (81.36) indicating a better HRQOL, and a lower mean score for HAQ-DI (0.37) indicating less activity limitation, compared with scores for the whole cohort at baseline (HAQ-DI 0.51; SF-36 PF10 75.91).

Table 4.12 GIS subscales, SF-36 PF10 and HAQ-DI scores of participants responding on different number of occasions

Baseline characteristic	All responders at **baseline** (n=1184)	Responded **only at baseline** (n=280)	Responders at baseline and at least **one other time-point** (n=904)	Responders at baseline and at least **two other time-points** (n=771)	Responders at baseline and at least **three other time-points** (n=655)	Responders at baseline and at least **all time-points** (n=510)
GIS Concern overall	48.65(28.33)	49.61(28.61)	48.37(28.26)	48.31(28.35)	47.86(28.25)	45.43(27.36)
GIS Medication side effects	40.45(26.33)	39.37(27.07)	40.77(26.12)	40.77(26.39)	39.97(25.77)	38.98(25.71)
GIS Unmet treatment need	33.46(20.57)	38.27(21.10)	32.07(20.22)	31.23(19.75)	30.42(19.61)	30.05(19.40)
GIS Wellbeing during an attack	45.19(26.41)	43.05(27.18)	45.83(26.16)	45.76(26.18)	45.33(26.22)	44.53(25.99)
GIS Concern during an attack	40.13(24.35)	39.7(25.14)	40.26(24.13)	39.85(23.99)	39.02(23.86)	38.23(23.64)
SF-36 PF10	75.91(26.12)	67.98(29.41)	78.34(24.54)	79.97(23.71)	80.84(23.30)	81.36(22.93)
HAQ-DI	0.51(0.71)	0.65(0.80)	0.46(0.68)	0.43(0.64)	0.41(0.625)	0.37(0.59)

Values are mean (SD)

Each **GIS subscale** scored from **0 to 100;** higher scores on each scale indicating a greater impact of gout on HRQOL/worse HRQOL.

SF-36 PF10 scored from **0 to 100**; higher score indicating performs all types of physical activities including the most vigorous without limitations due to health.

HAQ-DI scored from **0 to 3**; higher score indicating greater activity limitation.

Baseline data for cohort have been published previously (Chandratre et al 2018).

Socio-demographic characteristics

Participants who responded only at baseline were less likely to be male, married or cohabiting, have attended further education, or consume alcohol daily and more likely to be older, or more deprived, in comparison with participants who responded at more time-points (Table 4.13).

BMI and alcohol frequency

Participants who responded only at baseline were less likely to consume alcohol daily, in comparison with participants who responded at more time-points (Table 4.14).

No marked differences were observed in the BMI of participants who responded on a different number of occasions.

Table 4.13 Socio-demographic characteristics of participants responding on different number of occasions

Baseline characteristic	All responders at **baseline** (n=1184)	Responded **only at baseline** (n=280)	Responders at baseline and at least **one other time-point** (n=904)	Responders at baseline and at least **two other time-points** (n=771)	Responders at baseline and at least **three other time-points** (n=655)	Responders at baseline and **all time- points** (n=510)
Sex n(%)						
Male	990(83.6)	213(76.1)	777(86.0)	672(87.2)	579(88.4)	456(89.4)
Age mean (SD)	65.61(12.49)	67.41(13.07)	65.05(12.26)	65.08(11.70)	65.14(11.42)	65.41(10.90)
Age categories n(%)						
<40	26(2.2)	4(1.4)	22(2.4)	17(2.2)	14(2.1)	10(2.0)
40-49.9	116(9.8)	30(10.7)	86(9.5)	66(8.6)	52(7.9)	34(6.7)
50-59.9	210(17.7)	43(15.4)	167(18.5)	149(19.3)	127(19.4)	98(19.2)
60-69.9	343(29.0)	63(22.5)	280(31.0)	247(32.0)	214(32.7)	177(34.7)
70-79.9	339(28.6)	85(30.4)	254(28.1)	219(28.4)	188(28.7)	148(29)
>80	150(12.7)	55(19.6)	95(10.5)	73(9.5)	60(9.2)	43(8.4)
Neighbour deprivation status† n(%)						
Most deprived	369(31.2)	113(40.4)	256(28.3)	217(28.1)	181(27.6)	140(27.5)
Middle	405(34.2)	85(30.4)	320(35.4)	281(36.4)	244(37.3)	193(37.8)
Least deprived	410(34.6)	82(29.3)	328(36.3)	273(35.4)	230(35.1)	177(34.7)
Ethnic origin n(%)						
White UK/ European	1126(95.1)	258(92.1)	868(96.0)	746(96.8)	632(96.5)	500(98.0)
Asian	16(1.4)	4(1.4)	12(1.3)	9(1.2)	8(1.2)	4(0.8)
Afro Caribbean	2(0.2)	0(0)	2(0.2)	1(0.1)	1(0.2)	1(0.2)
African	2(0.2)	1(0.4)	1(0.1)	0(0)	0(0)	0(0)
Chinese	1(0.1)	1(0.4)	0(0)	0(0)	0(0)	0(0)
Other	7(0.6)	1(0.4)	6(0.7)	4(0.5)	4(0.6)	1(0.2)

† calculated using tertiles of indices of multiple deprivation

Table 4.13 cont. Socio-demographic characteristics of participants responding on different number of occasions

Baseline characteristic	All responders at **baseline** (n=1184)	Responded **only at baseline** (n=280)	Responders at baseline and at least **one other time-point** (n=904)	Responders at baseline and at least **two other time-points** (n=771)	Responders at baseline and at least **three other time-points** (n=655)	Responders at baseline and **all time-points** (n=510)
Relationship status n(%)						
Married	809(68.4)	183(65.4)	626(69.2)	540(70.0)	462(70.5)	360(70.6)
Widowed	114(9.6)	34(12.1)	80(8.8)	68(8.8)	54(8.2)	39(7.6)
Cohabiting	73(6.2)	11(3.9)	62(6.9)	48(6.2)	37(5.6)	29(5.7)
Divorced	69(5.8)	18(6.4)	51(5.6)	43(5.6)	37(5.6)	30(5.9)
Separated	22(1.9)	4(1.4)	18(2.0)	12(1.6)	11(1.7)	7(1.4)
Single	78(6.6)	21(7.5)	57(6.3)	51(6.7)	45(6.9)	39(7.6)
Attendance at further education n(%)	249(21.0)	42(15.0)	207(22.9)	186(24.1)	166(25.3)	136(26.7)

Baseline data for cohort have been published previously (Chandratre et al 2018).

Table 4.14 BMI and alcohol frequency of participants responding on different number of occasions

Baseline characteristic	All responders at **baseline** (n=1184)	Responded **only at baseline** (n=280)	Responders at baseline and at least **one other time-point** (n=904)	Responders at baseline and at least **two other time-points** (n=771)	Responders at baseline and at least **three other time-points** (n=655)	Responders at baseline and **all time-points** (n=510)
BMI kg/m^2 mean (SD)	29.13(5.11)	28.65(4.95)	29.27(5.15)	29.31(5.17)	29.17(5.12)	29.08(4.92)
BMI categories n(%)						
<18.5 kg/m^2	1(0.1)	0(0)	1(0.1)	1(0.1)	1(0.2)	1(0.2)
18.5-24.9 kg/m^2	219(18.5)	60(21.4)	159(17.6)	136(17.6)	120(18.2)	91(17.8)
25-29.9 kg/m^2	511(43.2)	111(39.6)	400(44.2)	340(44.1)	288(44.0)	228(44.7)
30-34.9 kg/m^2	260(22.0)	56(20.0)	204(22.6)	176(22.8)	147(22.5)	123(24.1)
35-39.9 kg/m^2	90(7.6)	21(7.5)	69(7.6)	59(7.7)	48(7.3)	33(6.5)
>40 kg/m^2	37(3.1)	5(1.8)	32(3.5)	29(3.8)	22(3.4)	16(3.1)
Alcohol frequency n(%)						
Never	113(9.5)	32(11.4)	81(9.0)	62(8.0)	53(8.1)	37(7.3)
Special occasions	155(13.1)	45(16.1)	110(12.2)	90(11.7)	73(11.1)	54(10.6)
1 to 3 times per month	109(9.2)	22(7.9)	87(9.6)	73(9.5)	62(9.5)	53(10.4)
Once or twice per week	254(21.5)	55(19.6)	199(22.0)	168(21.8)	140(21.4)	109(21.4)
3 to 4 times per week	263(22.2)	68(24.3)	195(21.6)	176(22.8)	149(22.7)	118(23.1)
Daily or almost daily	273(23.1)	49(17.5)	224(24.8)	196(25.4)	172(26.3)	136(26.7)

Baseline data for cohort have been published previously (Chandratre et al 2018).

4.4.3 Assessment of Item non-response

GIS subscale item non-response

The percentage of item non-response (percentage of responders at each time-point who did not provide a response to an individual item) for the items of the GIS subscale was below 10% for all items except two items from the GIS WBDA relating to missing work because of gout symptoms (proportion missing at different time-points ranged from 10.9 to 13.6%) and have difficulty working because of gout symptoms (9.4 to 12.1%) (Table 4.15). Other items with between 5 and 10% item non-response at any time-point included all of the items from the GIS MSE and GIS UTN subscale and selected items from the GIS WBDA subscale. Neither the GIS CO nor GIS CDA subscales had more than 5% item non-response at any time-point. With the exception of the GIS UTN subscale, less than 5% of responders had insufficient item responses to enable calculation of each GIS subscale total at each time-point (Table 4.18). 5% to 6.4% of responders had insufficient item responses to the GIS UTN subscale over the five time-points, to permit calculation of a GIS UTN score.

SF-36 PF 10 item non-response

The percentage of item non-response was below 5% for all items of the SF-36 PF10 except the item relating to vigorous activity at 12-month time-point where 5.4% was missing (Table 4.16). Over the five time-points, 18.9 to 21% of participants had insufficient item response to enable calculation of a SF-36 PF10 score (see Table 4.19).

HAQ-DI item non-response

The percentage of item non-response was below 5% for all items of the HAQ-DI except the item relating to taking a bath at baseline and 12 months which had 5.2% and 5.4% missing (Table 4.17). Over all five time-points fewer than 5% of participants had insufficient item responses to enable the calculation of a HAQ-DI score (see Table 4.19).

Table 4.15 Number and percentage of responders at each time-point who did not provide a response to each individual item of GIS questionnaire (item non-response to GIS subscale questionnaire items)

GIS item	Baseline n(%) missing	6 months n(%) missing	12 months n(%) missing	24 months n(%) missing	36 months n(%) missing
Gout Concern overall (CO)					
1A I am worried that I will have a gout attack within the next year	48(4.1)	21(2.6)	17(2.4)	20(2.9)	18(3.0)
1B I am afraid that my gout will get worse with time	41(3.5)	18(2.2)	20(2.8)	20(2.9)	19(3.1)
1C I worry that I will not be able to continue my leisure activities as a result of my gout	44(3.7)	23(2.8)	20(2.8)	23(3.3)	18(3.0)
1D I feel anxious that my gout will interfere with my future activities	45(3.8)	22(2.7)	19(2.6)	23(3.3)	17(2.8)
Gout Medication Side-effects (MSE)					
1E I am bothered by side-effects from my gout medications	63(5.3)	37(4.5)	33(4.6)	33(4.7)	34(5.6)
1K I worry about the long-term side-effects of gout medications	68(5.7)	42(5.1)	35(4.9)	32(4.6)	31(5.1)
Gout Unmet treatment need (UTN)					
1I My current medications are effective for treating a gout attack when I have one	79(6.7)	47(5.7)	39(5.4)	44(6.3)	33(5.5)
1L My current medications do not work well to prevent attacks from happening	106(9.0)	63(7.7)	53(7.4)	61(8.8)	42(6.9)
IM I have control over my gout	72(6.1)	32(3.9)	22(3.1)	28(4.0)	26(4.3)
Wellbeing during attack (WBDA)					
2 During your last gout attack, how much of the time did you experience the following?					
A Miss work because of your gout symptoms	161(13.6)	107(13.1)	93(12.9)	93(13.4)	66(10.9)
B Have difficulty working because of your gout symptoms	136(11.5)	99(12.1)	84(11.7)	82(11.8)	57(9.4)
C Have difficulty with recreational or social activities because of your gout symptoms	40(3.4)	31(3.8)	28(3.9)	35(5.0)	21(3.5)
D Have difficulty with self-care activities, e.g. feeding, bathing or dressing, because of your gout symptoms	33(2.8)	26(3.2)	25(3.5)	28(4.0)	19(3.1)
3 During your last gout attack, how much did your symptoms interfere with the following things?					
A Your mood	43(3.6)	29(2.6)	36(5.0)	30(4.3)	22(3.6)
B Your ability to move about	33(2.8)	21(2.6)	38(5.3)	28(4.0)	20(3.3)

C Your sleep	39(3.3)	30(3.7)	41(5.7)	31(4.5)	23(3.8)
D Your normal work	67(5.7)	54(6.6)	52(7.2)	45(6.5)	26(4.3)
E Your recreational activities	51(4.3)	31(3.8)	42(5.8)	35(5.0)	20(3.3)
F Your enjoyment of life	37(3.1)	28(3.4)	38(5.3)	29(4.2)	20(3.3)
G Your ability to do what you want to do	31(2.6)	25(3.1)	34(4.7)	28(4.0)	18(3.0)
Gout Concern during attack (CDA)					
1F I am mad or angry when I experience a gout attack	56(4.7)	26(3.2)	27(3.7)	24(3.4)	18(3.0)
1G It is difficult to plan ahead for events or activities because I may have a gout attack	46(3.9)	23(2.8)	23(3.2)	20(2.9)	19(3.1)
1H I feel depressed when I experience a gout attack	53(4.5)	22(2.7)	27(3.7)	23(3.3)	17(2.8)
1J I miss planned or important activities when I have a gout attack	53(4.5)	28(3.4)	21(2.9)	30(4.3)	20(3.3)

Table 4.16 Number and percentage of participants at each time-point who did not provide a response to each individual item of SF-36 PF10 questionnaire (item non-response to SF-36 PF10 questionnaire items)

SF-36 PF10 item	**Baseline** n(%) missing	**6 months** n(%) missing	**12 months** n(%) missing	**24 months** n(%) missing	**36 months** n(%) missing
Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports	44(3.7)	24(2.9)	39(5.4)	26(3.7)	18(3.0)
Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling or playing golf	29(2.4)	17(2.1)	29(4.0)	17(2.4)	13(2.1)
Lifting or carrying groceries	29(2.4)	17(2.1)	30(4.2)	17(2.4)	11(1.8)
Climbing several flights of stairs	36(3.0)	28(3.4)	30(4.2)	13(1.9)	15(2.5)
Climbing one flight of stairs	34(2.9)	18(2.2)	29(4.0)	14(2.0)	15(2.5)
Bending, kneeling or stooping	29(2.4)	22(2.7)	35(4.9)	16(2.3)	13(2.1)
Walking more than a mile	35(3.0)	21(2.6)	31(4.3)	18(2.6)	17(2.8)
Walking half a mile	38(3.2)	23(2.8)	32(4.4)	21(3.0)	13(2.10
Walking one hundred yards	31(2.6)	20(2.4)	28(3.9)	15(2.2)	16(2.6)
Bathing and dressing yourself	26(2.2)	14(1.7)	28(3.9)	12(1.7)	11(1.8)

Table 4.17 Number and percentage of participants at each time-point who did not provide a response to each individual item of HAQ-DI questionnaire (item non-response to HAQ-DI questionnaire items)

HAQ-DI item	**Baseline** n(%) missing	**6 months** n(%) missing	**12 months** n(%) missing	**24 months** n(%) missing	**36 months** n(%) missing
Dress					
A Dress yourself include tying shoe-laces and doing buttons	41(3.5)	15(1.8)	21(2.9)	9(1.3)	16(2.6)
B Shampoo your hair	52(4.4)	14(1.7)	24(3.3)	12(1.7)	15(2.5)
Arise					
C Stand from a chair	46(3.9)	17(2.1)	27(3.7)	10(1.4)	16(2.6)
D Get in and out of bed	50(4.2)	11(1.3)	25(3.5)	8(1.1)	16(2.6)
Eat					
E Cut your meat	49(4.1)	15(1.8)	24(3.3)	10(1.4)	15(2.5)
F Open a milk carton	43(3.6)	15(1.8)	25(3.5)	8(1.1)	14(2.3)
G Lift a full glass or cup to your mouth	42(3.5)	12(1.5)	24(3.3)	10(1.4)	14(2.3)
Walk					
H Walk outdoors on flat ground	40(3.4)	15(1.8)	24(3.3)	9(1.3)	15(2.5)
I Climb up 5 steps	41(3.5)	14(1.7)	24(3.3)	11(1.6)	14(2.3)
Hygiene					
J Wash and dry your entire body	43(3.6)	14(1.7)	22(3.1)	8(1.1)	13(2.1)
K Take a tub bath	61(5.2)	33(4.0)	39(5.4)	22(3.2)	27(4.5)
L Get on and off the toilet	40(3.4)	13(1.6)	22(3.1)	9(1.3)	13(2.1)
Reach					
M Reach and get a 5 lb object from above your head	40(3.4)	11(1.3)	23(3.2)	8(1.1)	14(2.3)
P Bend down and pick up clothing from the floor	42(3.5)	15(1.8)	25(3.5)	8(1.1)	16(2.6)
Grip					
N Open car doors	40(3.4)	15(1.8)	25(3.5)	11(1.6)	14(2.3)
O Open jars that have been previously opened	42(3.5)	13(1.6)	24(3.3)	8(1.1)	14(2.3)
Q Turn taps on and off	42(3.5)	15(1.8)	24(3.3)	10(1.4)	15(2.5)
Daily					
R Run errand and shop	42(3.5)	20(2.4)	27(3.7)	13(1.9)	19(3.1)
S Get in and out of a car	39(3.3)	17(2.1)	25(3.5)	12(1.7)	19(3.1)
T Do vacuuming or yard work	37(3.1)	17(2.1)	30(4.2)	13(1.9)	15(2.5)

Table 4.18 Proportion of responders at each time-point with insufficient item responses to enable calculation of each GIS subscale total

GIS subscale	**Baseline** n(%) missing scores	**6 months** n(%) missing scores	**12 months** n(%) missing scores	**24 months** n(%) missing scores	**36 months** n(%) missing scores
GIS Concern overall (CO)	38(3.2)	17(2.1)	18(2.5)	19(2.7)	17(2.8)
GIS Medication side-effects (MSE)	54(4.6)	29(3.5)	26(3.6)	23(3.3)	27(4.5)
GIS Unmet treatment need (UTN)	76(6.4)	42(5.1)	36(5.0)	40(5.7)	32(5.3)
GIS Wellbeing during attack (WBDA)	34(2.9)	26(3.2)	35(4.9)	29(4.2)	19(3.1)
GIS Concern during attack (CDA)	45(3.8)	19(2.3)	22(3.1)	19(2.7)	17(2.8)

Table 4.19 Proportion of responders with insufficient item responses to enable calculation of SF-36 PF10 and HAQ-DI

	Baseline n(%) missing scores	**6 months** n(%) missing scores	**12 months** n(%) missing scores	**24 months** n(%) missing scores	**36 months** n(%) missing scores
SF-36 PF10	224(18.9)	152(18.6)	142(19.7)	146(21.0)	123(20.3)
HAQ-DI	40(3.4)	11(1.3)	22(3.1)	8(1.1)	13(2.1)

Gout-specific variables, comorbidities, socio-demographic variables, BMI and alcohol missing data

Variables with greater than 10% missing data included serum urate level at 61.1%, eGFR 21.4%, body pain at 18.2%, and PHQ-9 score at 12% (Table 4.20). Variables with between 5 and 10% missing data included number of gout flares (but less than 5% missing after baseline), reported oligo/polyarticular flares (greater than 5% only at 36 months), disease duration, tophi, self-reported allopurinol use, prescription for allopurinol, prescription for NSAIDs, GAD-7 score, BMI and attendance at further education. The missing data for tophi and prescription for allopurinol were attributable to participants who did not consent to medical note review (n=105). All other variables had less than 5% missing data.

Table 4.20 Missing data for gout-specific variables, comorbidities, socio-demographic variables, BMI and alcohol frequency

	Baseline n(%) missing	**6 months** n(%) missing	**12 months** n(%) missing	**24 months** n(%) missing	**36 months** n(%) missing
Gout Specific					
Number of flares[†]	61(5.2)	17(2.1)	20(2.8)	31(4.5)	25(4.1)
Current flare	49(4.1)	12(1.5)	20(2.8)	30(4.3)	21(3.5)
Oligo/polyarticular flares	53(4.5)	13(1.6)	21(2.9)	31(4.5)	31(5.1)
Disease duration	89(7.5)				
Record of tophi[◇]	105(8.9)				
Serum urate Level[◇]	723(61.1)				
Allopurinol use*	64(5.4)	38(4.6)	29(4.0)	34(4.9)	31(5.1)
Allopurinol prescription[◇]	105(8.9)				
Comorbidities					
Diabetes Mellitus	0(0)				
CVA	0(0)				
Hypertension	0(0)				
TIA	0(0)				
Hyperlipidaemia	0(0)				
MI	0(0)				
Angina	0(0)				
Renal Failure	0(0)				
Renal Calculi	0(0)				
Total number comorbidities	0(0)				
eGFR <60	253(21.4)				
NRS pain	63(5.3)	44(5.4)	29(4.0)	35(5.0)	27(4.5)
Body pain	215(18.2)				
Depression (PHQ 9 score)	142(12.0)		41(5.7)		41(6.8)
Anxiety (GAD 7 score)	90(7.6)		36(5.0)		23(3.8)
Socio-demographic					
Sex	0(0)				
Age	0(0)				
Ethnicity	30(2.5)				
IMD	0(0)				
Marital status	19(1.6)				
Attendance at further education	65(5.5)				
BMI & alcohol					
BMI	65(5.5)	16(2.0)	14(1.9)	23(3.3)	10(1.7)
Alcohol	17(1.4)				

Grey shading = variable was not collected at this time-point

[†] **Number of gout flares** in previous 12 months at baseline, 12 months and 36 months, in previous 6 months at 6 and 12 months; *** Self-reported allopurinol** use taken from questionnaire responses[◇] **In medical record** in the two years prior to baseline; **total number of comorbidities** self-reported in baseline questionnaire (diabetes, hypertension, hyperlipidaemia, myocardial infarction, angina, cerebrovascular accident, transient ischaemic attack, renal failure, renal calculi); **eGFR <60** mL/min/1.73m^2 indicative of chronic kidney disease; **NRS pain in last week** ranges from 0 (no pain) to 10 (pain as bad as it can be); **body pain** (including ache or discomfort or stiffness) for one day or longer in the 4 weeks prior to baseline**; PHQ-9 score** ranges from 0 to 27 Minimal depression 0-4, Mild 5-9, Moderate 10-14, Moderately Severe 15-19, Severe 20-27; **GAD-7 score** ranges from 0 to 21 Minimal anxiety 0-4, Mild 5-9, Moderate 10-14, Severe 15-21; **IMD** indices of multiple deprivation; **BMI** body mass index kg/m^2

4.5 Discussion

The aim of this chapter was to describe the cohort that will be analysed in chapters five, six, seven and eight and to investigate potential sources of bias within the cohort. The flow of participants through the cohort was described and the potential for attrition bias and item non-response missing data was investigated. The detailed reporting of an observational cohort study undertaken in this chapter, including participant flow, assessment of potential sources of bias such as attrition bias, and the extent of missing data for each exposure of interest, has been advocated by the 'strengthening the reporting of observational studies in epidemiology' (STROBE) guidelines (Vandenbroucke et al 2007).

4.5.1 Flow of participants through the cohort study

The adjusted response to the questionnaire at baseline was 65.9% of eligible mailed participants, with subsequent adjusted responses of 79.1%, 72.8%, 75.4% and 68.4% at 6, 12, 24 and 36 months respectively. Whilst there is debate regarding what constitutes an acceptable response in epidemiological studies (Morton et al 2012), a response of 60% or more has been advocated as acceptable (Bowling, 2009; Fincham, 2008; Sim & Wright, 2000). Thus, an acceptable adjusted response was observed at every time-point. However, it is important to acknowledge that attrition of participants was observed during follow-up, due to either exclusion from mailing and refusal or non-response. Various reasons for attrition were identified including deaths, address or contact problems, health reasons, participants not wishing to take part, or returning a questionnaire too late. Attrition is a common occurrence in prospective cohort studies (Bankhead, Aronson & Nunan, 2018; Lacey, Jordan & Croft, 2013), particularly in cohorts of older persons (Lacey, Jordan & Croft, 2013). Several strategies detailed in section 3.7, were deployed in this study to maximise response and minimise bias in this cohort.

Of the 1184 participants who responded at baseline, 605 (51%) responded to the questionnaire at 36 months. The attrition observed is lower than that reported in prospective cohorts in people with gout by Becker et al (attrition 53% over one year) (2009) and Khanna et al (74% over two years) (2011a). However, lower attrition rates were reported by Dalbeth et al (2013) where of the 291 participants who attended at baseline, 249 (85.6%) provided information at one year via a mailed survey. The difference in the attrition observed in the cohort analysed in this chapter, in comparison with these published studies, could be attributable to this cohort being based in primary care, the sole use of a questionnaire survey (Dalbeth et al 2013 and Khanna et al 2011a utilised clinic visits in their studies), a longer follow-up period, or the older age of this cohort. Comparable attrition has been observed in a prospective cohort study postal questionnaire survey of back pain in primary care in the West Midlands; Lacey, Jordan & Croft (2013) reported that of 11,209 respondents at baseline, 5,121 went on to respond at three years (46%) thus demonstrating a similar overall proportion of attrition to those described in this chapter.

4.5.2 Assessment of attrition bias and comparison with existing studies

Whilst the attrition of participants from a cohort study is common, attrition bias is not an inevitable consequence of attrition (Lacey, Jordan & Croft, 2013). Thus, the potential for attrition bias, where attrition caused the characteristics of responders to be systematically different than those who did not participate (Chatfield, Brayne & Matthews, 2005; Kirstman, Manno & Cote, 2005), was assessed in this chapter.

Responders in this cohort study tended to have a slightly longer disease duration and less frequent gout flares, compared with non-participants. A higher proportion of responders had a serum urate greater than 360 µmol/L, although the responders were more likely to have a record of their serum urate in their medical records. More of the responders also reported taking allopurinol or had a prescription for either allopurinol or NSAIDs in their medical record.

Responders were less likely to report a history of diabetes, MI or angina, and had slightly lower PHQ-9, GAD-7, pain NRS, and global health scores, compared with non-participants. Responders also had lower GIS UTN scores, higher SF-36 PF10 scores, and lower HAQ-DI scores, indicating better HRQOL and less activity limitation respectively. Responders were less likely to be over the age of 80 years, female or classified as most deprived, but more reported attending further education, or consuming alcohol daily or almost daily. However, many of the differences observed in this chapter between responders and non-participants were small.

A limited number of studies have compared the characteristics of responders and non-participants in prospective cohort studies of gout patients. In contrast to this cohort, Becker et al (2009) reported that participants who did not complete follow-up in a prospective cohort study in treatment failure gout had more swollen and tender joints at baseline compared with those who remained in the study. There were no differences between completers and non-completers in gout flare frequency, disease duration, or serum urate levels. In addition, Stewart et al (2018) reported no difference in flare frequency, disease duration or activity limitation at baseline between participants in a prospective cohort study of gout patients at one year compared with those lost to follow-up. However, these smaller studies included participants recruited from secondary care thus with different gout severity in comparison to this primary care cohort. In addition, existing studies have failed to compare a wide range of characteristics when comparing attrition bias. These findings in this chapter are consistent with the observations in the same cohort at baseline reported by Roddy et al (2015) who found that responders at baseline were more likely to be male and were less socioeconomically deprived. However, responders at baseline tended to be older compared with non-participants, whilst this investigation has shown that older participants were more likely to be lost to follow-up after baseline.

The findings of this prospective cohort are also consistent with other studies and multivariable systematic reviews investigating attrition bias. A shorter disease duration, higher modified-HAQ, and depression scores for participants who dropped out of a longitudinal study in Rheumatoid Arthritis has been described by Iannaccone et al (2013). In an analysis of attrition in cohort studies in primary care Lacey, Jordan & Croft (2013) also reported a lower prevalence of depression in responders and reported that responders in prospective questionnaire surveys had a higher number of prescriptions for disease specific medications. In cohorts of older participants, attrition has been shown to be more likely in participants with low levels of physical activity or poor functioning in multivariate analysis (Brilleman, Pachana & Dobson, 2010; Chatfield, Brayne & Matthews, 2005). Other studies have also reported that participants who self-reported poor health status are more likely to be lost to follow-up (Brilleman, Pachana & Dobson, 2010; Schmidt et al 2011).

Chatfield, Brayne & Matthews (2005) described how older age was associated with attrition in a multivariable analysis. Potential factors contributing to more attrition post baseline for older participants may have been cognitive impairment, frailty or death, which are more common in older persons and are factors acknowledged to be associated with a greater chance of attrition (Brilleman, Pachana & Dobson, 2010; Chatfield, Brayne & Matthews, 2005). Other reported risk factors for attrition include being female, living alone or being unmarried (Chatfield, Brayne & Matthews, 2005) and a lower educational attainment (Brilleman, Pachana & Dobson, 2010; Chatfield, Brayne & Matthews, 2005; Gustavson et al 2012). Conversely, Galea & Tracey (2007) stated that a higher socio-economic status has been associated with increased participation in studies.

The analysis undertaken in this chapter has provided a greater understanding of the characteristics of participants who responded to this survey. These findings are helpful when

considering the generalisability of the findings in this thesis to the wider population of people living with gout who may be encountered in clinical practice.

The group of participants responding at all five time-points had a longer disease duration, more allopurinol use, less frequent gout flares, lower PHQ-9, GAD-7, pain NRS, global health, GIS UTN, HAQ-DI scores and higher SF-36 PF10 scores. These finding suggest that the inclusion of only participants who responded at each time-point could potentially yield a cohort with less generalisable characteristics compared to all responders. As these differences were in key dependent and independent outcomes this could have influenced results if only complete cases (participants who responded at all five time-points) had been used in the analyses in this thesis. This justifies the inclusion where possible of all responders when analysing outcome measures in this thesis.

4.5.3 Item non-response missing data

Missing data is a common occurrence in quantitative research (Dong & Peng, 2013; Peyre, Leplege & Coste, 2011) and has the potential to lead to bias (Biering, Hjollund & Frydenberg, 2015; Bland, 2015), a reduction in sample sizes and thus potentially misleading results (Biering, Hjollund & Frydenberg, 2015; Fayers & Machin, 2016; Peyre, Leplege & Coste, 2011). When evaluating missing data there is no universally accepted level of missing data which is considered acceptable (Dong & Peng, 2013). It has been proposed that less than 5% missing data may lead to less biased inferences (Dong & Peng, 2013; Schafer, 1999) whilst an alternative level of less than 10% missing has been described as being less likely to yield biased results from statistical analysis (Bennett, 2001).

The analysis in this chapter revealed that less than 5.2 % of self-reported gout flare data was missing at any time-point, this is a dependent variable in the analysis undertaken in the next two chapters (five and six) where further analysis and discussion of missing data for this variable can be found.

Most of the GIS subscales, SF-36 PF10, and HAQ-DI items had less than 5% missing responses, however there were some exceptions including items related to work on the GIS WBDA subscale, items related to medication on the GIS UTN & MSE subscales, the item related to vigorous activities on the SF-36 PF10, and the item related to taking a bath on the HAQ-DI. The questionnaires did not allow participants to indicate where items were not applicable. In this older cohort, questions relating to employment may have been less relevant to some participants. When one considers that not all participants had a prescription for Allopurinol within their medical record, and the widely reported suboptimal prescribing of medication to manage gout symptoms (Kuo et al 2015a), it is possible that many participants may have felt that questions about gout medication were less relevant to them and thus did not respond to these items. Higher proportions of missing data for the vigorous activity item of the SF-36 PF10 have been reported previously by older participants, where many participants reported they felt the question was irrelevant as they were unable to do any vigorous activity (Hayes et al 1995). It is also possible that the proportion of missing data on the item of the HAQ-DI relating to taking a bath was higher because those participants without a bath felt the question was less relevant to them.

Missing data from item non-response may be particularly problematic if the missing data means there are insufficient responses to permit calculation of a total score (Peyre, Leplege & Coste, 2011). Except for the UTN subscale (which had 5 to 6.4% missing), all the GIS subscales had less than 5% of total scores for participants missing. However, the SF-36 PF10 had more than 10% of total scores for participants missing, due to insufficient items to calculate the total score. The HAQ-DI had less than 5% of total scores for participants missing. These HRQOL measures are the dependent variables which were statistically modelled in chapter eight where further discussion of missing data for these outcomes can be found.

The percentage of missing data at each time-point was below 10% for all the gout-specific, comorbid, socio-demographic and other variables, apart from data relating to serum urate level, eGFR, body pain, and PHQ-9 score. The very high percentage of missing data relating to serum urate levels is perhaps not surprising when one considers that suboptimal monitoring of serum urate has been reported in other cohorts of gout patients (Wall et al 2010) and the acknowledged suboptimal management of gout in the wider population (Kuo et al 2015a). The percentage of participants with missing eGFR data in this cohort is very similar to the proportion of patients registered with GPs in the West Midlands missing a recent eGFR from their primary care records reported by Cottrell, Chambers & O'Connell (2012). Over 10% missing data for the PHQ-9 has also been reported in other cohorts living with chronic disease (Wang et al 2016), and over 10% of pain data was found to be incomplete in a postal questionnaire investigating multiple body site pain and HRQOL in people living with osteoarthritis in primary care in the West Midlands (Lacey et al 2014).

Variables with 5 to 10% of missing data at any time-point included a history of oligo/polyarticular flares, disease duration, tophi, allopurinol use, GAD-7, BMI, and attendance at further education. Some missing data for variables obtained from medical record is attributable to not all participants consenting to medical record review.

As some missing data was identified in the main outcome measures to be modelled in this thesis, gout flares (chapter five & six) and HRQOL measures GIS subscales, SF-36 PF10 and HAQ-DI (chapters seven & eight), these findings had implications for the type of statistical modelling techniques which can be used. Consequently, statistical modelling techniques which can accommodate missing data in longitudinal dependent variables were used to analyse gout flares (latent class growth analysis) and GIS subscales, SF-36 PF10 and HAQ-DI (linear mixed models) (Berlin, Parra & Williams, 2014; Cheng et al 2010; Jung & Wickrama, 2008; Locascio & Atri, 2011).

4.5.4 Strengths and limitations

This chapter follows good practice recommended when reporting observational studies, by describing the participants in this prospective study, reasons for non-participation, the proportion of missing data, and the potential sources of bias (Vandenbroucke et al 2007).

As attrition can arise because participants may respond intermittently to a survey (Cumming & Golstein, 2016; Twiske & de vente, 2002; Twisk, 2003) or because participants may drop out completely from a survey (Twiske & de vente, 2002; Twisk, 2003), the potential for attrition bias was assessed in this chapter using different approaches. A wide range of characteristics were compared including gout-specific characteristics, comorbidities, socio-demographic characteristics and HRQOL scores. Some differences between responders versus non-participants, and also between participants who responded only at baseline compared with participants who responded at multiple time-points, were identified. A decision was made not to undertake any statistical analysis of the differences between the different groups as, due to the wide variety of characteristics compared, this would have involved multiple comparisons and multiple statistical tests. Such multiple statistical testing has the potential to lead to type 1 error; suggesting there was a difference between groups when in fact there was not (Sedgwick, 2014). In addition, undertaking statistical testing on differences which were not an a priori hypothesis has been highlighted as poor practice when investigating the consequences of attrition (Dumville, Torgerson & Hewitt, 2006).

The investigation of the proportion of missing data in independent and dependent variables was important due to the potential consequences for sample sizes, interpretation of results and the potential for bias (Biering, Hjollund & Frydenberg, 2015; Fayers & Machin, 2016). However, it is important to acknowledge that the mechanism of missing data can be as important as the proportion of missing data (Dong & Peng, 2013). Additional analysis could have been undertaken to investigate the potential mechanism of missing data, whether the

missing data was missing completely at random (MCAR), missing at random (MAR) or missing not at random (MNAR). Such tests could have included Little's missing completely at random (MCAR) test, or statistical test comparing the difference between groups with complete and missing data, or undertaking logistic regression (Dong & Peng, 2013). However, tests such as Little's MCAR test cannot indicate whether the mechanism is MAR, but instead provides evidence for or against the mechanism being MCAR (Fielding, Fayers & Ramsay, 2010). It is not possible to distinguish between whether data is MAR and MNAR based on observed data, as the values of unobserved data are not known (Fielding, Fayers & Ramsay, 2009; Kristman, Manno & Cote, 2005).

Whilst various sources of potential bias were investigated with in this chapter, it is important to acknowledge that there is still the potential that other sources of bias may exist with in this cohort. Further potential sources of bias are discussed in chapter nine. However, in a prospective study of HRQOL measures of this nature, it is rarely possible to identify all potential sources of bias (Fayers and Machin, 2016).

4.6 Conclusion

In conclusion, this chapter described the flow of participants through this cohort study and investigated attrition bias and item non-response missing data. Some small differences were observed between the characteristics of participants who responded at different time-points compared with non-participants. The proportion of independent and dependent variable missing data was below 10% for the majority of variables. The findings of this chapter were used to inform the analysis in future chapters by influencing how gout flares (chapter five and chapter six) and HRQOL measures (chapter eight) were statistically modelled. The aim of the next chapter is to describe the frequency and pattern of self-reported gout flares in this prospective cohort of people living with gout in primary care.

5 Chapter Five Frequency and pattern of gout flares

5.1 Overview of chapter and aim

The previous chapter described the cohort of participants in this study and investigated potential sources of bias within the cohort. The aim of this chapter is to describe the frequency and pattern of self-reported gout flares in this prospective cohort of people living with gout in primary care.

5.2 Background

5.2.1 Latent trajectory modelling; a patient-centred approach

The within-person pattern of a longitudinal outcome over time is referred to as a trajectory (Curran, Obeidat & Losardo, 2010; Nagin & Odgers, 2010). Individuals with outcomes which follow a similar within-person pattern over time can be grouped together, to create distinct groups of individual trajectories which were previously unobserved or latent (Nagin, 1999; Nagin & Land, 1993). Research methods which are focused on detecting groups of individuals based on similarity in latent trajectories have been described as person-centred approaches; as the focus is on the relationship between individuals, rather than the relationship between variables (Muthen & Muthen, 2000). Such methods include latent class growth analysis and growth mixture modelling (Muthen & Muthen, 2000) and have the potential to detect inter-individual differences in intra-individual patterns of change over time (Curran, Obeidat & Losardo, 2010; Jung & Wickrama, 2008). This is in contrast with a variable centred approach where the focus would be on the relationship solely among variables (Muthen & Muthen, 2000).

5.2.2 Latent growth curve models

An initial step in identifying latent trajectories in cohorts involves latent growth curve modelling (LGCM) (Berlin, Parra & Williams, 2014), which can also be referred to as growth modelling, latent growth modelling, latent curve analysis, growth curve analysis, latent variable growth curve modelling, latent trajectory modelling and latent trajectory analysis (Burant, 2016; Berrington, Smith & Sturgis, 2006; Curran & Hussong, 2003; Duncan & Duncan, 2009). During such modelling a growth curve is specified which best describes the intra-individual change in observed longitudinal outcomes for a whole cohort (Stoel, van de Wittenoer & Hox, 2004; Wang & Bodner, 2007). The shape of the growth curve describing the intra-individual change can assume any number of polynomial forms such as linear, quadratic or piecewise (Wickrama et al 2016).

5.2.3 Assumptions and limitations of latent growth curve models

LGCMs require data from a minimum of three, ideally more, time-points with the same measurement instrument used on each occasion (Curran & Hussong, 2003; Duncan & Duncan, 2009). An advantage of current approaches to undertaking latent growth curve modelling is that not all individuals are required to have repeated measures data recorded at all time-points, thus permitting the inclusion of individuals with partially missing data (Curran, Obeidat & Losardo, 2010; Hox & Stoel, 2005). It has been proposed that inclusion of individuals with partially missing data is associated with less bias, compared with the inclusion of individuals with only complete data (Curran & Hussong, 2003). Current approaches to undertaking latent growth curve modelling also afford greater flexibility regarding the time points at which repeated measures are recorded and thus unequally spaced time points are permitted (Curran, Obeidt & Losardo, 2010). A further advantage of current approaches to undertaking latent growth curve modelling is that methods of estimation exist which do not assume that residuals of repeated measures are normally distributed (Curran & Hussong, 2003).

Specifying a single LGCM to describe the change in longitudinal data, follows the assumption that all individuals come from a single population and that a single latent growth curve can adequately describe growth parameters for the entire population (Jung & Wickrama, 2008; Ram & Grimm, 2009). This limitation is of particular relevance when modelling clinical outcomes in health research as change in symptoms is likely to vary amongst individuals within a cohort (Nagin & Odgers, 2010; Wickrama et al 2016). Despite this limitation, it is important to specify a LGCM prior to progressing on to more complicated models, as determining the functional form of the latent growth curve will inform the type of models used in further analysis (Curran & Hussong, 2003; Curran, Obeidat & Losardo, 2010). Latent growth curve modelling was undertaken in this chapter, in order to inform the latent class growth analysis undertaken in the next chapter (chapter six).

5.2.4 Latent class growth analysis

Latent class growth analysis (LCGA) is a person-centred method which can be considered as a special type of latent variable modelling (Berlin, Parra & Williams, 2014; Jung & Wickrama, 2008; Muthen & Muthen, 2000). LCGA models allocate individuals to different groups or classes based on the shape of their latent growth curve trajectory (Muthen & Muthen 2000). Thus, each class is summarised by latent growth curve with growth parameters. For example, in a quadratic model the growth parameters are a mean intercept, a mean linear slope and a mean quadratic slope. The class's intercept and slopes are referred to as latent parameters as these parameters were unobserved prior to undertaking the analysis (Wickrama et al 2016). In LCGA, the variance and covariance within each latent class is eliminated by fixing the variance of the intercept and slope(s) to zero (Berlin, Parra & Williams, 2014; Jung & Wickrama, 2008; Muthen & Muthen, 2000; Wickrama et al 2016). Thus, due to the lack of within-class variance, LCGA models assume that all individual growth trajectories within a specific class are homogenous (Berlin, Parra & Williams, 2014; Jung & Wickrama,

2008;Wickrama et al 2016). LCGA can thus be thought of as a fixed effect model (Peugh & Fan, 2012). For example, for a quadratic model all class members have the same intercept, linear slope and quadratic slope (Berlin, Parra & Williams, 2014; Wickrama et al 2016). Individuals are probabilistically assigned to the latent class which best reflects their latent trajectory; with individuals assigned to the class for which they have the highest posterior probability (Andruff et al 2009). The LCGA process sequentially increases the number of latent classes, until the optimal number of classes is determined (Wickrama et al 2016).

A limitation of LCGA is that the lack of within-class variance may be unrealistic and it has the potential to generate more classes than a model which permits within class variance (Wickrama et al 2016). Widely acknowledged advantages to using LCGA include the reduced computational burden and faster model conversion (Wickrama et al 2016). Thus, LCGA is advocated as a method to determine latent classes prior to progressing to growth mixture modelling (Berlin, Parra & Williams, 2014; Jung & Wickrama, 2008).

5.2.5 Growth mixture modelling

Growth mixture modelling (GMM) is also a longitudinal modelling technique which estimates a different latent growth curve for each latent class within a population (Muthen & Muthen, 2000; Jung & Wickrama, 2008; Peugh & Fan, 2012). The LGCMs for each class have mean intercepts and slope(s) but also, unlike LCGA, (Berlin, Parra & Williams, 2014; Jung & Wickrama, 2008), GMM permits variance and covariance of the intercept and slope (Berlin, Parra & Williams, 2014; Jung & Wickrama, 2008; Wang & Bodner, 2007). Thus, GMM can be considered a random effects model (Peugh & Fan, 2012). The mixture of distributions of observed variables which are permitted within this type of model is acknowledged by the inclusion of ‘mixture’ in the name growth mixture modelling (Peugh & Fan, 2012; Wickrama et al 2016).

Whilst the advantage of GMM is the ability to accommodate within class variability, this method also has limitations (Ram & Grimm, 2009; Wang & Bodner, 2007). GMM requires the estimation of more latent parameters than LCGA, which requires an increased computational load and increases with the number of latent classes in the model (Wang & Bodner, 2007). This has the potential to lead to less stable models with convergence and computational problems (Wang & Bodner, 2007).

5.3 Method and Analysis Plan

5.3.1 Aim and Objectives

Aim

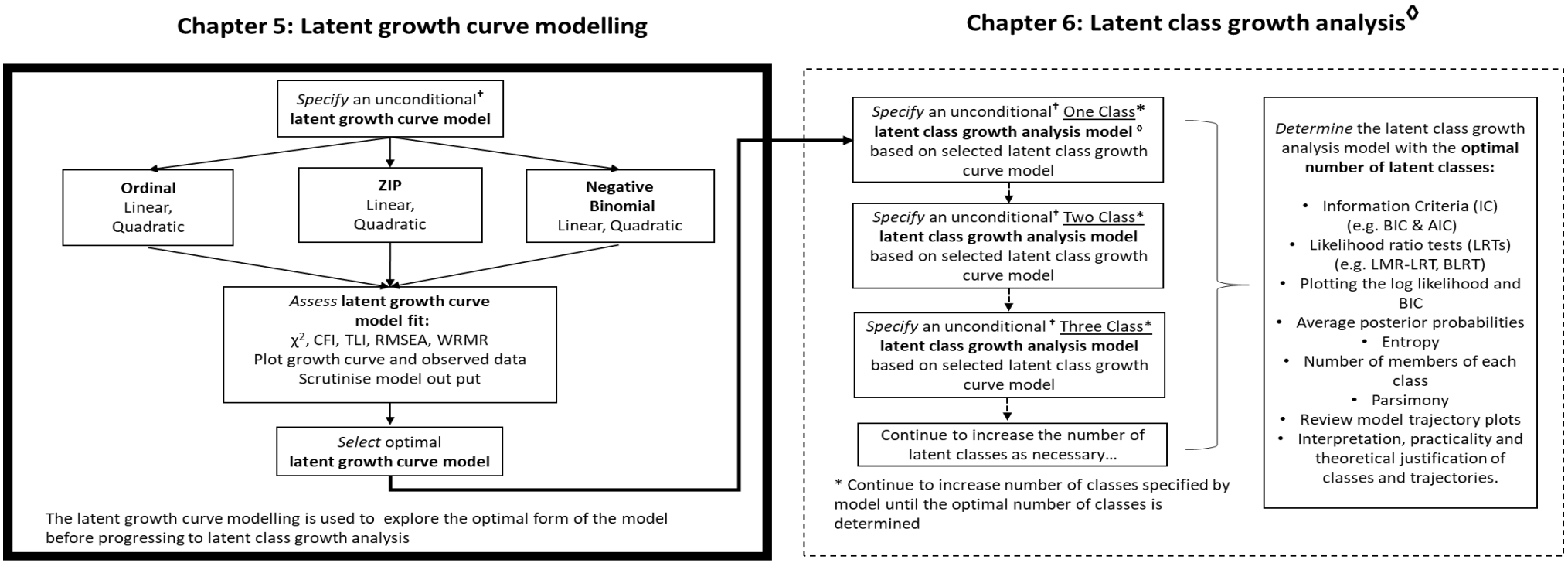
The aim of this chapter is to describe the frequency and pattern of self-reported gout flares in a prospective cohort of people living with gout in primary care.

Objectives

The objectives of this chapter are to:

1. Use descriptive statistics and spaghetti plots to describe the frequency and pattern of self-reported gout flares.
2. Use latent growth curve modelling (LGCM) to determine the shape of the growth curve which best describes the longitudinal gout flare data.

Figure 5.1 on the following page is a flowchart summary of the stages of the latent growth curve modelling used in this chapter and the latent class growth analysis used in chapter 6



◊ In latent class growth analysis the variance of intercepts and slopes are fixed to zero. Growth mixture modelling (where intercept and slope variance is permitted) was also undertaken following the same stages. †unconditional model without covariates

AIC Akaike Information Criteria; **BIC** Bayesian information criteria; **BLRT** Bootstrap likelihood ratio test; **CFI** Comparative Fit Index; **LMR-LRT** Lo Mendell Rubin likelihood ratio test; **TLI** Tucker-Lewis index; **RMSEA** Root mean square error of approximation; **WRMR** Weighted root-mean-square residual; **χ2** Chi-Squared Statistic; **ZIP** zero-inflated Poisson

Figure 5.1 Flowchart summary of latent growth curve modelling and latent class growth analysis

5.3.2 Data source

The data for this secondary analysis were derived from the study described in chapter three. The primary observed variable used for the analysis in this chapter was the number of gout flares reported by participants at each questionnaire time-point (baseline, 6, 12, 24 and 36 months). The number of gout flares in the previous 12 months was reported in the baseline, 24, and 36-month questionnaires, whilst the number of flares reported in the previous 6 months was reported in the 6, and 12-month questionnaires.

5.3.3 Ordinal data considerations

The gout flare data were considered to be ordinal data, as respondents were asked to report the number of gout flares they experienced in a specified time period by selecting one of the response options 0, 1, 2, 3, 4, ≥5.

Ordinal data are frequently modelled as continuous data in research (Feldman, Masyn & Conger, 2009). If there are a large number of categories and the distribution of responses are symmetrical then modelling ordinal data as continuous may not distort estimates or model fit (Hox & Stoel, 2005). However, where there is a smaller number of categories, ordinal data should be modelled in a manner which takes the ordinal nature of the data in to account, as treating ordinal data as continuous data has the potential to lead to biased estimates and errors in fit statistics (Feldman, Masyn & Conger, 2009). Also ordinal data are often skewed, having a greater proportion of responses in the lowest category, which can make growth modelling with ordinal data a challenge, and require that the data are modelled appropriately (Feldman, Masyn & Conger, 2009).

Thus, within this analysis the gout flare data were coded specifically as categorical ordinal data. Such ordinal data can be thought of as observed response variables in categories, which

are actually taken from a range of latent unobserved continuous response variables (Curran & Hussong, 2003; Hox & Stoel, 2005).

5.3.4 Count data considerations

The number of self-reported gout flares can also be considered to be count data. Consequently, in addition to undertaking the latent growth curve modelling which treated the gout flare data as ordinal data, the data were also modelled treating the data as count data.

Poisson

Count data, which follow a Poisson distribution, are particularly common in health research but are frequently modelled incorrectly (Shiyko, Yuelin & Rindskopf, 2012). Where latent growth curve modelling and latent class growth analysis use count data following a Poisson distribution, the observed variable Y is not modelled directly instead the natural log of the Poisson parameter is used e.g. $\ln(\lambda_{it})$ (Shiyko, Yuelin & Rindskopf, 2012), where λ_i represents the mean and variance parameter of the Poisson distribution and the equivalence of the mean and the variance is a familiar characteristic of the Poisson distribution (Liu & Powers, 2007).

Zero-inflated Poisson

Not all count data follow a Poisson distribution (Bohning et al 1999; Lambert, 1992; Liu & Powers, 2007). Medical and health research frequently yield count data which contain many zeros and if the prevalence of zeros is not considered when modelling data there is the potential for biased results and inferences (Lee et al 2006; Yoon et al 2015). Where count data include an excessive proportion of zeros an alternative distribution, the zero-inflated Poisson (ZIP) distribution, is recommended (Lambert, 1992). Unlike the Poisson distribution, which is defined by just one parameter (λ_i) the ZIP distribution is defined by two parameters λ_i and p_i (which is the probability of being zero-inflated) (Liu & Powers, 2007).

Negative binomial

The Poisson distribution and Poisson models assume equidispersion; where the mean and the variance of count data are equal (Hilbe, 2007; Zhou & Carin, 2015). However, count data are often overdispersed; where the variance of the count data is greater than the mean (Hilbe, 2007; Zhou & Carin, 2015). Such overdispersion has the potential to affect the reliability of assessments of model fit and the model parameter estimates (Hilbe, 2007). Where there is overdispersion in count data, the assumptions of Poisson models are not valid and negative binomial models are recommended for the latent variable modelling of over dispersed count data (Reincke & Sedding, 2011; Zhou & Carin, 2015).

The potential for zero-inflation or overdispersion of the gout flare data modelled in this thesis was assessed via the methods described in section 5.3.5.

5.3.5 Prior to modelling data

Where possible, the use of prior theory or existing research has been recommended prior to undertaking latent variable modelling; to guide a hypothesis relating to the potential functional form of latent growth curves and also the number of anticipated classes that may be found (Belin, Parra & Williams, 2013; Ram & Grimm, 2009). However, until this point there has not been any research undertaken relating specifically to latent variable modelling of gout flares.

In the absence of prior research, a preliminary stage in LGCM was the inspection of raw data to inform the modelling of the data (Berlin, Parra & Williams, 2014; Berrington, Smith & Sturgis, 2006). The proportion of participants reporting 0, 1, 2, 3, 4, or ≥ 5 or more flares at each time point was displayed in a stacked column chart. This descriptive analysis of the gout flare count data revealed the proportion of participants reporting zero flares. It is important to assess the proportion of zeros within this data, as the data could be considered to be count

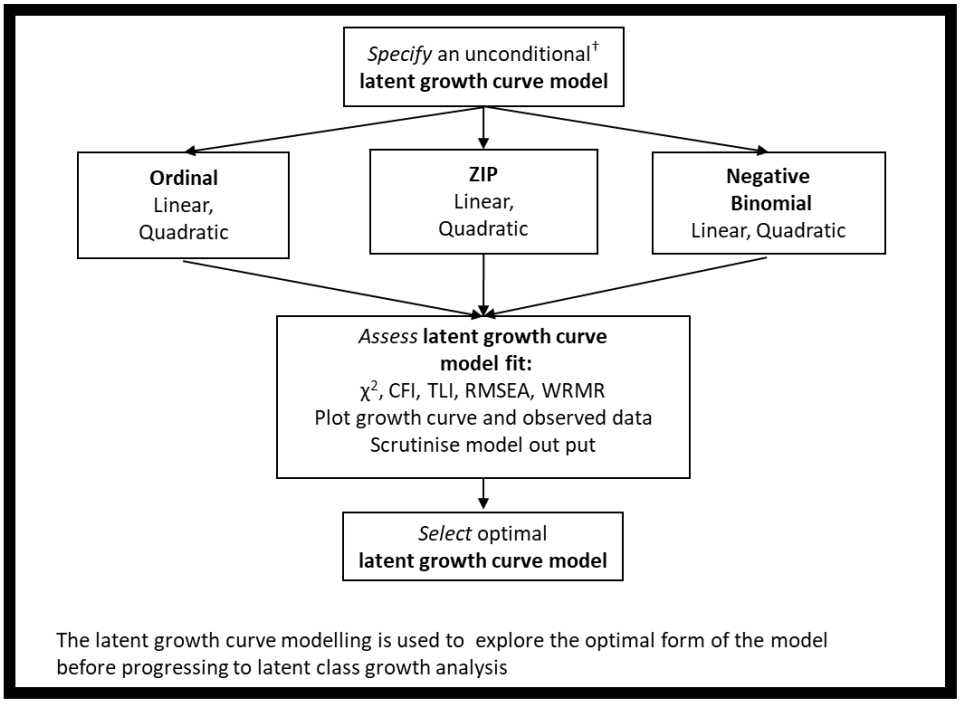
data and a higher proportion of zeros would indicate that the data is likely to follow a zero-inflated Poisson (ZIP) distribution, rather than a Poisson distribution (Lambert, 1992). At each time-point, the mean number of gout flares was compared to the variance of the gout flares, to test whether the mean was equal to variance; as would be the case with Poisson distributed data (Liu & Powers, 2007) but not the case with overdispersed data (Hilbe, 2007; Zhou & Carin, 2015). Any difference in distribution would need to be taken into consideration when modelling data (Bohning et al 1999; Lambert, 1992; Liu & Powers, 2007). The median (interquartile range, IQR) number of gout flares at each time-point was also calculated. The number of gout flares reported by each participant at each available time point was displayed in a spaghetti plot, where a single line linked repeated measures over time for each participant. An interpolation line was also added to the spaghetti plot, to demonstrate which functional form displayed the overall change in gout flares.

SPSS version 24 was used for the descriptive statistics and spaghetti plots.

5.3.6 Latent growth curve modelling

Specifying a single unconditional latent growth curve model

The first stage in latent variable modelling of the longitudinal data was to specify a baseline single class unconditional (with no covariates included) LGCM (Jung & Wickrama, 2008; Ram & Grimm 2009; Wickrama et al 2016). All participants who self-reported how many gout flares they had experienced on at least one time-point were included in the modelling. To determine the optimal LGCM for the data-set, a series of unconditional latent growth curves were fitted to the data (Ram & Grimm 2009), the types of latent growth curves models fitted can be found in Figure 5.2. These included linear and quadratic ordinal models, linear and quadratic zero-inflated Poisson (ZIP) models and linear and quadratic negative binomial (NB) models. M*plus* version 8.1 was used for the latent growth curve modelling.



†**unconditional model without covariates**
CFI Comparative Fit Index; **TLI** Tucker-Lewis index; **RMSEA** Root mean square error of approximation; **WRMR** Weighted root-mean-square residual; **χ2** Chi-Squared Statistic; **ZIP** zero-inflated Poisson

Figure 5.2 Stages of latent growth curve modelling undertaken in this chapter

Assessment of model fit of unconditional latent growth curve models

It was important to evaluate the fit of the different LGCMs generated in order to determine which type of LGCM would be used in the latent class growth analysis in the next chapter (Burant, 2016; Curran & Hussong, 2003; Currang, Obeidat & Losardo, 2010). Several factors need to be considered when evaluating the goodness of fit of LGCMs (Wickrama et al 2016). Consequently, a range of model fit indices were reviewed, the latent growth curve models were plotted, and the latent growth curve model results were inspected.

Review of model fit indices

The use of a range of model fit indices, rather than sole reliance on one model fit indices, has been recommended as different indices reflect different aspects of a latent growth curve model fit (Hooper, Coughlan & Mullen, 2008; Wu, West & Taylor, 2009).

Table 5.1 below displays the model fit indices that were considered in this chapter and how the results from the model input were interpreted.

Table 5.1 Model fit indices and interpretation of fit for latent growth curve modelling

Indices *	**Perfect fit**	**Good fit**	**Acceptable fit**
Chi-Squared Statistic (χ^2)	-	P value >0.05 (Hooper, Coughlan & Mullen, 2008; Wickrama et al 2016)	-
Comparative Fit Index (CFI) (Bentler, 1990)	1.0 (Wickrama et al 2016)	0.95 or more (Hu & Bentler, 1999)	0.90 or more (Hu & Bentler, 1999)
Tucker-Lewis Index (TLI) (Tucker & Lewis, 1973)	1.0 (Wickrama et al 2016)	0.95 or more (Hu & Bentler, 1999)	0.90 or more (Hu & Bentler, 1999)
Root mean square error of approximation (RMSEA) (Steiger & Lind, 1980)	0.0 (Wickrama et al 2016)	0.06 or less (Hu & Bentler, 1999)	0.08 or less (MacCallum, Browne & Sugawara, 1996)
Weighted root-mean-square residual (WRMR)	-	-	Around 1.10 or less for latent growth model with 5 timepoints and n= >500 (Yu, 2002)

*A description of each of these indices, along with the indices used in chapter six, can be found in appendix 10.

Plotting latent growth curve models

Each of the different latent growth curves were examined graphically by plotting the latent growth curve. This permitted a visual comparison of the growth curve of the actual longitudinal data, with the growth curve estimated by the latent growth curve model. Where the data were modelled as ordinal data, the probability for specific categories was calculated (Lee, Wickrama & O'Neal, 2017). For the ordinal data, the probability of reporting two or more gout flares at each time point was plotted. The probability of two or more gout flares was selected for the y axis of the plots, as when patients experience two or more flares in the last 12 months it has been advocated that urate-lowering therapy should be advised (Hui et al 2017), as highlighted in chapter one. For the count data, the mean number of gout flares for the estimated model and the sample were plotted.

5.3.7 Sensitivity analysis

All latent growth curve models were initially modelled including participants who had reported the number of gout flares they had experienced on at least one time-point. Latent growth curve models were also then specified using two other different data sets: i) those participants who had reported gout flares on at least three time-points and ii) those participants who had reported gout flares at every time-point.

5.4 Results

5.4.1 Gout flare data descriptive results

1164 participants self-reported gout flares on at least one time-point and were included in the analysis in this chapter. Twenty participants were excluded from this analysis as they did not report the number of gout flares experienced at any time-point.

At each time-point, the largest proportion of participants reported experiencing no gout flares; 35.4%, 56.8%, 57.8%, 51.6% and 54.7% at baseline, 6, 12, 24 and 36 months respectively

(Table 5.2 and Figure 5.3). Of the participants who reported experiencing gout flares more people reported experiencing one flare (ranged from 13.8-20.6% at different time-points), followed by two flares (9.3-16.7%), three flares (5.6-9.2%) and five or more flares (5.8-12.2%). The smallest proportion (2.6-6%) reported experiencing four gout flares. In the analysis sample, the number of participants with missing gout flare data due to no response to the gout flare frequency item in the questionnaire (item-non response) ranged from 1.5 to 3.5%, whilst the number of participants with missing gout flare data due to a questionnaire not being returned (unit-non response) ranged from 0 to 48.0%.

The mean number of gout flares was 1.66 baseline and 1.18 at 36 months, whilst the median was 1 at baseline and 0 at 36 months. The variance was 2.94 at baseline and 2.69 at 36 months.

Table 5.2 Self-reported gout flares at each time point for participants who reported at ≥1 time-points (n=1164)

	Baseline (n=1123)	**6 Months** (n=801)	**12 Months** (n=701)	**24 Months** (n=665)	**36 Months** (n=580)
0 flares n(%)	398(35.4)	455(56.8)	405(57.8)	343(51.6)	317(54.7)
1 flare n(%)	231(20.6)	127(15.9)	121(17.3)	107(16.1)	80(13.8)
2 flares n(%)	187(16.7)	86(10.7)	67(9.6)	62(9.3)	74(12.8)
3 flares n(%)	103(9.2)	45(5.6)	47(6.7)	52(7.8)	34(5.9)
4 flares n(%)	67(6.0)	21(2.6)	20(2.9)	33(5.0)	22(3.8)
≥5 flares n(%)	137(12.2)	67(8.4)	41(5.8)	68(10.2)	53(9.1)
Mean flares* (Standard Deviation)	1.66 (1.72)	1.06 (1.57)	0.97 (1.46)	1.29 (1.71)	1.18 (1.64)
Median flares* (IQR)	1.0 (0,3)	0.0 (0,2)	0.0 (0,1.5)	0.0 (0, 2)	0.0 (0,2)
Variance*	2.94	2.47	2.12	2.92	2.69
≥2 gout flares n(%)	494 (42.4)	219(27.3)	175(24.9)	215(32.3)	183(31.3)
Item response missing†	41(3.5)	17(1.5)	20(1.7)	31(2.7)	25(2.3)
Unit response missing†	0(0)	346(29.7)	443(38.1)	468(40.2)	559(48.0)

*calculated treating category ≥5 as 5 flares †denominator n=1164

Baseline gout flare data for cohort have been published previously (Chandratre et al 2018).

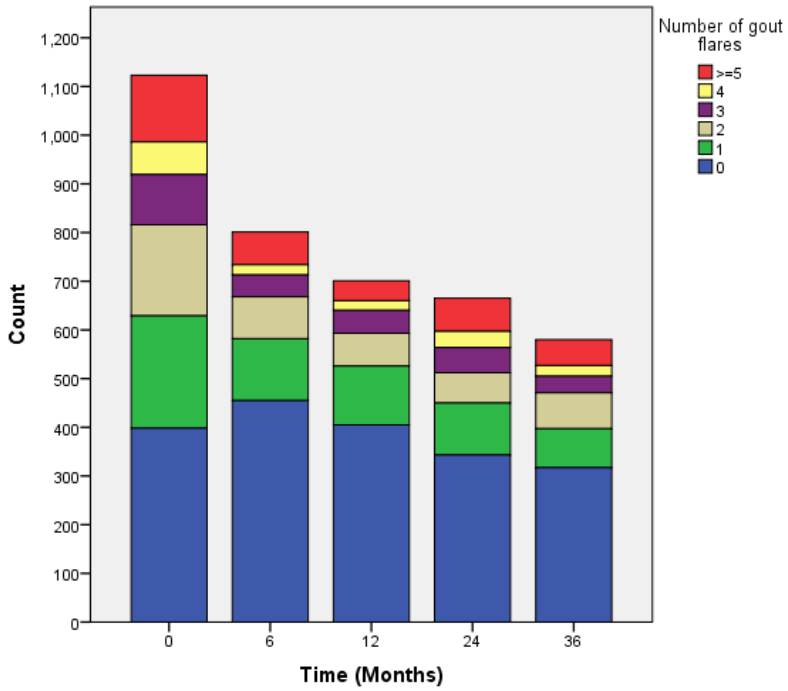


Figure 5.3 Proportion of participants reporting gout flares in each frequency category at each time-point

Baseline gout flare data for cohort have been published previously (Chandratre et al 2018).

5.4.2 Individual participant gout flare pattern of change

The spaghetti plot shown in Figure 5.4 plots the reported gout flares for each participant as a line. The lines in the figure show the diverse range of different patterns of reported gout flares over time. An interpolation line is shown as the bold black line in the dataset, suggested a quadratic functional form.

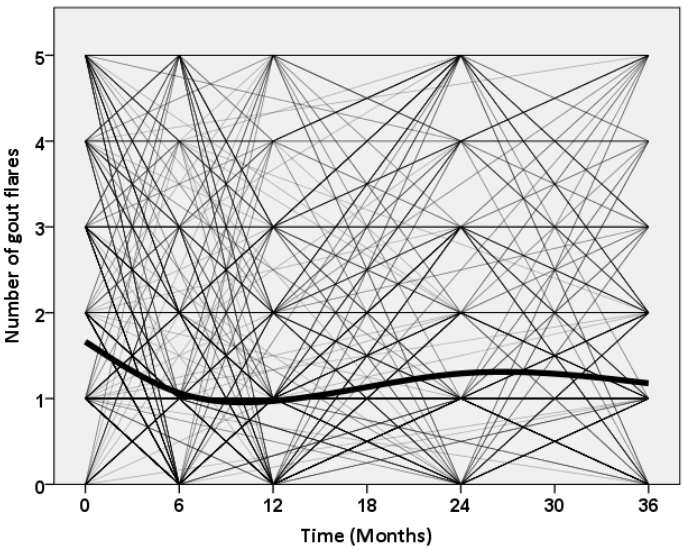


Figure 5.4 Spaghetti plot of self-reported gout flares for all participants self-reporting the frequency of gout flares at ≥1 time-points (n=1164)

5.4.3 Latent growth curve modelling

Ordinal latent growth curve models

The p-values for the chi-squared statistics were statistically significant for both the linear and the quadratic ordinal LGCM (Table 5.3). The CFI and TLI for the quadratic ordinal LGCM were 0.97 and 0.99 respectively, which were slightly higher than the CFI and TLI of 0.95 and 0.98 for the linear model. The RMSEA (with confidence intervals) for the quadratic LGCM of 0.08 (0.07,0.09) was lower than the RMSEA for the linear model which was 0.10 (0.09, 0.11). The WRMR result was also lower for the quadratic model, 1.71, than the result of 2.31 returned from the linear model.

Table 5.3 Latent growth curve model fit indices and results for linear and quadratic ordinal models

Models	χ^2	**CFI**	**TLI**	**RMSEA (CI)**	**WRMR**
Linear Ordinal LGCM	278.06 p<0.0001	0.95	0.98	0.10 (0.09, 0.11)	2.31
Quadratic Ordinal LGCM*	174.63 p<0.0001	0.97	0.99	0.08 (0.07, 0.09)	1.71

*fixing the quadratic slope to zero variance improved the running of the model

In the plot of the quadratic LGCM (see Figure 5.5), the data points in the plots for the observed data and estimated model were closer together at all time-points (except at 24 months), compared to the observed data and estimated model plots for the linear model (Figure 5.6).

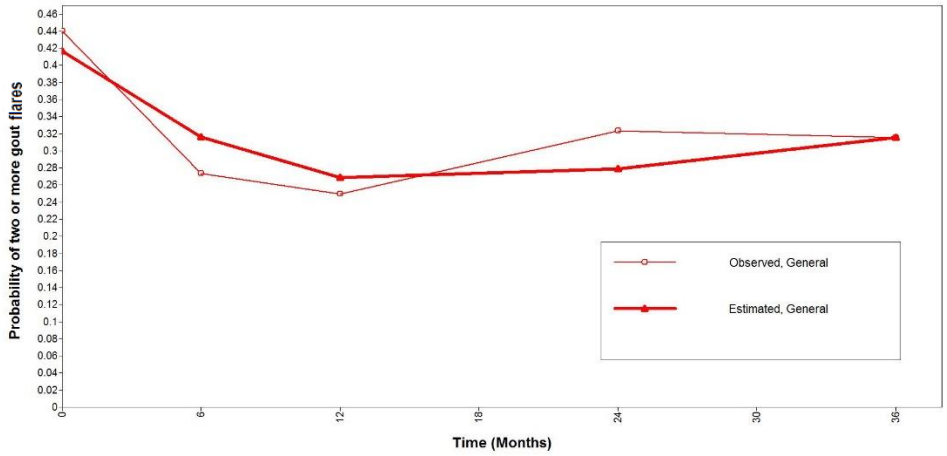


Figure 5.5 Quadratic ordinal latent growth curve model plot; the probability of ≥2 gout flares at each time-point

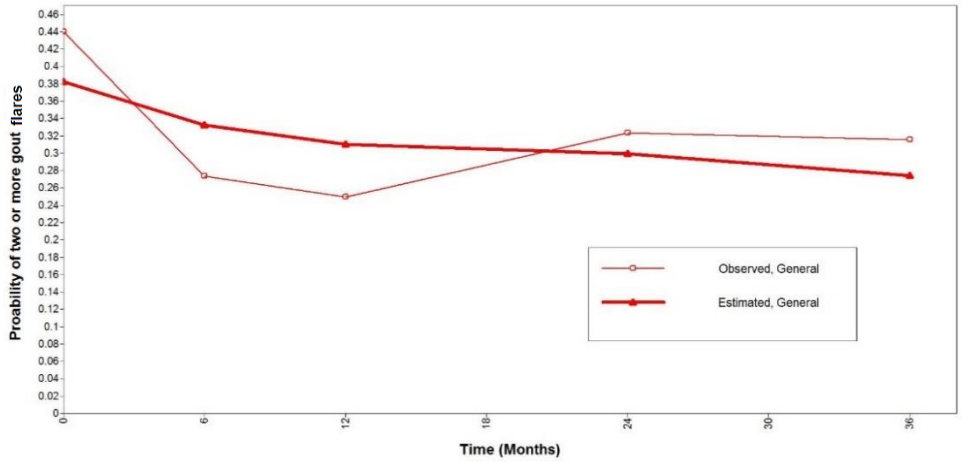


Figure 5.6 Linear ordinal latent growth curve model plot; the probability of ≥2 gout flares at each time-point

Zero-inflated Poisson and negative binomial latent growth curve models

The ZIP LGCMs required a change to the M*plus* coding to change the integration method, in order to successfully run. Yet despite these amendments the linear and quadratic ZIP LGCM output still displayed warnings that the model was modified to avoid singularity, warnings that some of the standard errors of the model parameter estimates were not trustworthy and two separate warnings of the possible non-identification of the model. These warnings remained despite amending the code in several ways including changing the number of integration points.

Both the linear and quadratic negative binomial models had a longer computational time. Despite amending the code and allowing both models to run over several days the models failed to run successfully and output from the models was not obtained.

The performance of the ZIP and negative binomial models did not improve when smaller data sets (those participants who had reported gout flares on at least three time-points (n=729) and those participants who had reported gout flares at every time-point (n=437)) were used.

5.4.4 Sensitivity analysis for latent growth curve modelling

In addition to undertaking the LGCM including participants who had reported the number of gout flares they had experienced on at least one time-point, latent growth curve models were also then specified using two other different data sets: i) those participants who had reported gout flares on at least three time-points (n=729) and ii) those participants who had reported gout flares at every time-point (n=437).

For those participants who had reported gout flares for at least three time-points the chi-squared statistics were statistically significant for both the linear and the quadratic ordinal LGCM (Table 5.4). The CFI and TLI for the quadratic ordinal LGCM were 0.98 and 0.99 respectively, which are slightly higher than the CFI and TLI of 0.96 and 0.98 for the linear model. The RMSEA for the quadratic LGCM of 0.09 (0.08, 0.11) was lower than the RMSEA for the linear model which was 0.11 (0.10, 0.13). The WRMR result was also lower for the quadratic model, 1.55, than the result returned from the linear model of 1.99.

In the plot of the quadratic LGCM (Figure 5.7), the data points in the plots for the observed data and estimated model are closer together at all time-points (except at 24 months), compared to the observed data and estimated model plots for the linear model (Figure 5.8).

Table 5.4 Latent growth curve model fit indices (for model including those participants who reported gout flares at ≥3 time-points)

Models	χ^2	CFI	TLI	RMSEA (CI)	WRMR
Linear Ordinal LGCM	226.07 p<0.0001	0.96	0.98	0.11 (0.10, 0.13)	1.99
Quadratic Ordinal LGCM*	151.71 p<0.0001	0.98	0.99	0.09 (0.08, 0.11)	1.55

*fixing the quadratic slope to zero variance improved the running of the model

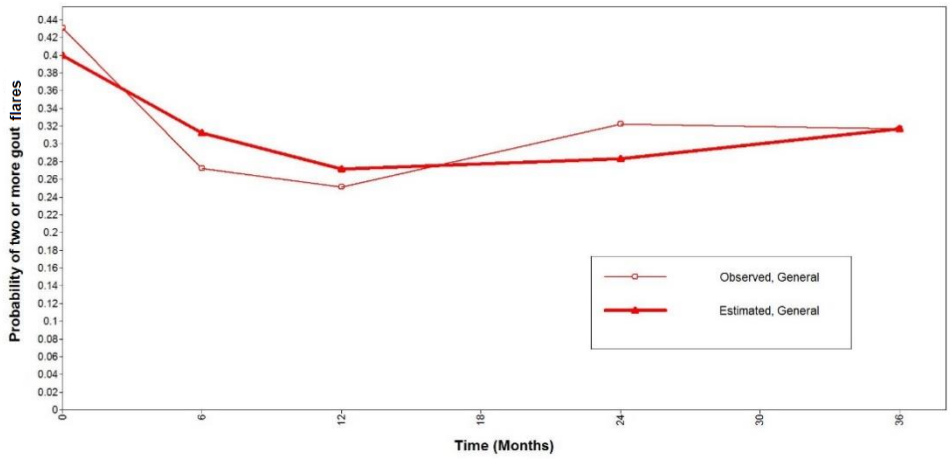


Figure 5.7 Quadratic ordinal latent growth curve model plot; the probability of ≥2 gout flares at each time-point (including those participants who had reported gout flares at ≥3 time-points)

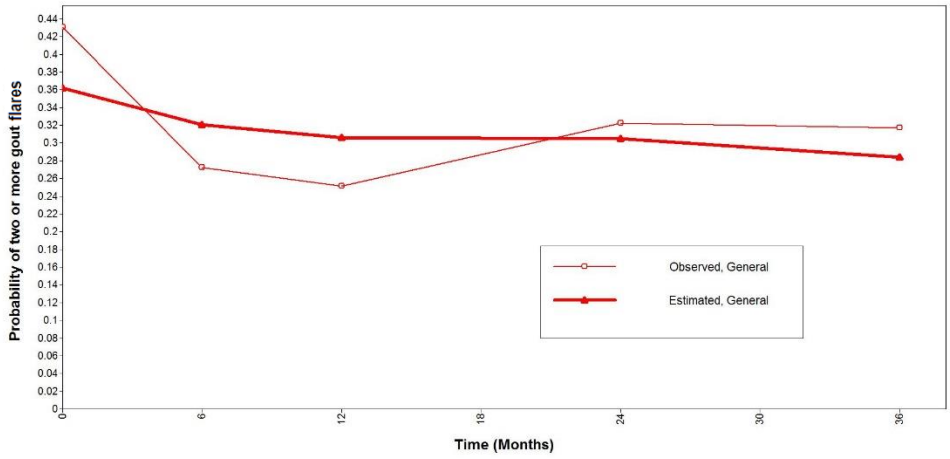


Figure 5.8 Linear ordinal latent growth curve model plot; the probability of ≥2 gout flares at each time-point (including those participants who had reported gout flares at ≥3 time-points)

For those participants who had reported gout flares at every time-point the chi-squared statistics were statistically significant for both the linear and the quadratic ordinal LGCM (see Table 5.5). The CFI and TLI for the quadratic ordinal LGCM were 0.98 and 0.99 respectively, which were slightly higher than the CFI and TLI of 0.97 and 0.98 for the linear model. The RMSEA for the quadratic LGCM of 0.08 (0.07, 0.10) was lower than the RMSEA for the linear model which was 0.10 (0.09, 0.12). The WRMR result was also lower for the quadratic model, 1.14, than the result returned from the linear model of 1.44.

In the plot of the quadratic LGCM (Figure 5.9), the data points in the plots for the observed data and estimated model were closer together at all time-points (except at 24 months), compared to the observed data and estimated model plots for the linear model (Figure 5.10).

Table 5.5 Latent growth curve model fit indices (for model including those participants who reported gout flares at all five time-points)

Models	χ^2	**CFI**	**TLI**	**RMSEA (CI)**	**WRMR**
Linear Ordinal LGCM	125.44 p<0.0001	0.97	0.98	0.10 (0.08, 0.12)	1.44
Quadratic Ordinal LGCM*	85.41 p<0.0001	0.98	0.99	0.08 (0.07, 0.10)	1.14

*fixing the quadratic slope to zero variance improved the running of the model

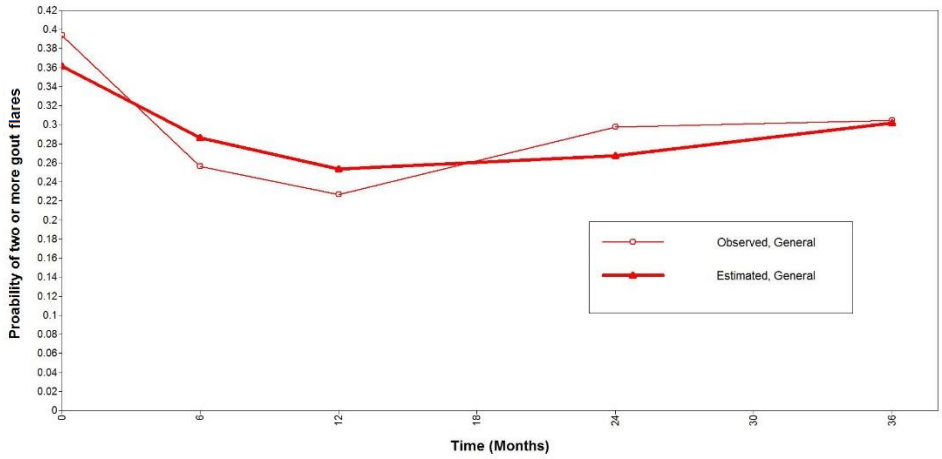


Figure 5.9 Quadratic ordinal latent growth curve model plot; the probability of ≥2 gout flares at each time-point (including those participants who had reported gout flares at every time-point)

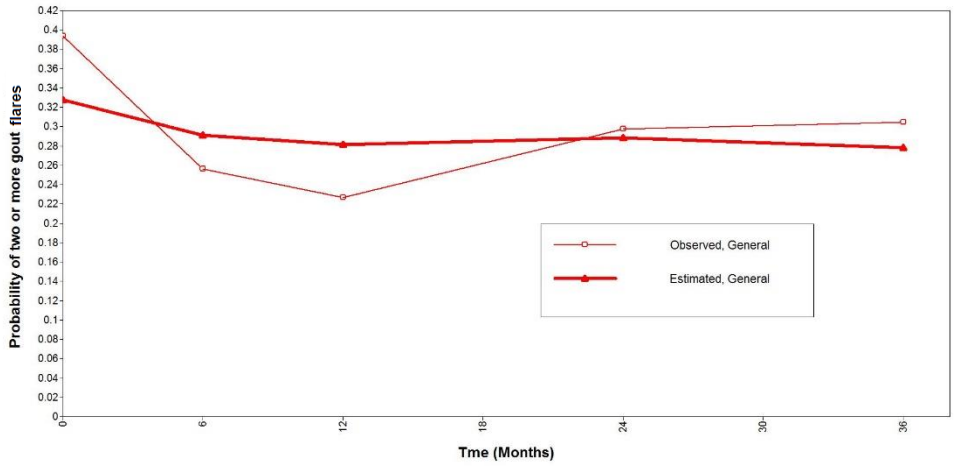


Figure 5.10 Linear ordinal latent growth curve model plot; the probability of ≥2 gout flares at each time-point (including those participants who had reported gout flares at every time-point)

5.5 Discussion

5.5.1 Gout flares

This chapter describes the frequency of gout flares self-reported by participants at five time-points in a prospective cohort study. At each time-point, the largest proportion of participants reported experiencing no gout flares in the reference period. The median number of gout

flares was 1 at baseline and 0 at 36 months, whilst the mean was 1.66 at baseline and 1.18 at 36 months.

Previous studies where patients have self-reported gout flare frequency over a 12-month period have reported a higher mean number of gout flares of 3.1-3.9 (ten Klooster et al 2014; Spaetgens et al 2015) and a median number of gout flares of 1 (IQR 0,3) (Scire et al 2013). In a cross-sectional survey by Khanna et al (2012c), 23% of participants reported experiencing no gout flares over 12 months, which is lower than the proportion reporting no flares within 12 months within this cohort. The lower gout flare frequency observed in this cohort compared to other studies could reflect the primary care setting, whereas ten Klooster et al (2014), Scire et al (2013) and Spaetgens et al (2015) all recruited from secondary care Rheumatology clinics and Khanna et al (2012c) used online recruitment. A lower mean number of flares of 0.51 was reported by Saseen et al (2012) in a subgroup of patients pre-defined by having 'infrequent' gout flares. Conversely, a higher mean number of flares of 6.4-6.8 was reported by Strand et al (2012) but this was in a cohort of patients with refractory chronic gout thus these results are less likely to be generalizable to a primary care gout population.

Despite the low mean and median gout flares, there was still a substantial proportion (24.9-42.4%) of the cohort who reported two or more gout flares at each time-point. This finding is comparable with the 25% of participants, with self-reported physician diagnosed gout, who reported two or more gout flares in the population-based survey by Proudman et al (2019). Frequent gout flares, as reported by more than a quarter of this cohort at each time-point, is associated with poorer HRQOL (Hirsch et al 2008; Khanna et al 2012c; Scire et al 2013), functional disability (Lee et al 2009; Scire et al 2013), higher use of healthcare resources (Jackson et al 2015; Khanna et al 2012c; Saseen et al 2012), and lower productivity (Khanna et al 2012c). Experiencing two or more gout flares in the previous 12 months has been highlighted as a situation where ULT should be particularly advised in order to optimally

manage gout symptoms (Hui et al 2017). The identification of a substantial proportion of participants within this cohort reporting frequent gout flares is consistent with previous observations that gout is frequently poorly managed (Kuo et al 2015a).

The spaghetti plot of gout flares self-reported at each time-point show participants reporting a diverse range of different patterns in gout flares over time. The pattern of gout flares over-time was investigated further through latent growth curve modelling as discussed in section 5.5.2.

5.5.2 Latent growth curve modelling

Ordinal models

The results of the latent growth curve modelling indicate a closer fit for the quadratic ordinal latent growth curve model, in comparison to other models trialled. The CFI and TLI results returned for both the linear and quadratic ordinal LGCM suggest good fit for both models as the results were greater than the level of 0.95 suggested by Hu & Bentler (1999) to indicate good fit. However, the results returned by the quadratic model were higher than those results returned from the linear model, indicating a potentially better fit. The RMSEA of 0.08 for the quadratic model indicates an acceptable fit as it was within the criteria, of 0.08 or less, suggested by MacCallum, Browne & Sugawara (1996). The RMSEA for the linear model of 0.10 (0.09, 0.11) suggested a less favourable fit for the model as the result, along with the confidence intervals, was above 0.08. The chi-squared result for both the quadratic and linear model was statistically significant. However, the study sample size was 1164 and a statistically significant result is often returned for the chi-squared statistic where there is a large sample size in latent growth curve modelling (Bentler & Bonnet, 1980). The WRMR result returned for the quadratic model (1.71) was lower than the result returned for the linear model (2.31) and a smaller WRMR had been proposed to suggest better fit (Wu, West & Taylor, 2009) and

the quadratic model returned a result which was closer to the value suggested by Yu (2012) as suggesting acceptable fit.

In addition to the model fit indices results obtained, the spaghetti plots and LGCM plots suggest that the gout flare data appear to display more of a quadratic form rather than a linear form. In the plot of the quadratic LGCM the data points in the plots for the observed data and estimated model are closer together at all time-points (except at 24 months), compared to the observed data and estimated model plots for the linear model; thus suggesting that the quadratic model displays a closer fit compared with the observed data.

Sensitivity analysis

The linear and quadratic ordinal LCGM were run using 3 different data sets; i) participants who self-reported gout flares on at least one time-point (n=1164), ii) those who had reported gout flares on at least three time-points (n=729) and iii) those who had reported gout flares at every time-point (n=437). The model fit indices results displayed a similar pattern in each data set; with a slightly higher CFI and TLI but lower RMSEA and WRMR returned for the quadratic ordinal LGCM. The same pattern in the plots of the linear and quadratic ordinal LGCM in the three different data sets were also observed. Thus, the results from the sensitivity analysis showed that a quadratic model was favoured in all three data sets.

Zero-inflated Poisson and negative binomial models.

The descriptive analysis prior to latent growth curve modelling, displayed zero-inflation in the self-reported gout flares; as the majority of participants at every time point reported no gout flares. This zero-inflation in the data and the fact that the gout flares data might be considered 'count data' justified modelling a zero-inflated Poission latent growth curve model (Liu & Powers, 2007).

The variance of the gout flare data was greater than the mean number of gout flares at each time-point. This finding indicated over dispersion in the data, thus providing justification for modelling a negative binomial latent growth curve model (Zhou & Carin, 2015).

However, when the ZIP and negative binomial models were run several problems were encountered; including multiple warnings in the output of the ZIP and excessive computational time with the negative binomial model. Such problems were not encountered when running the ordinal latent growth curve models.

5.5.3 Strengths and limitations

Gout flares

A strength of this study was the analysis of gout flare data from five time-points enabling the investigation of the dynamic nature of the change in gout flares overtime. Flares have been identified as a key reported outcome measure by Outcome Measures in Rheumatology Clinical Trials (OMERACT) (Schumacher et al 2009). However, a standardised definition of a gout flare outcome measure has yet to be endorsed as a clinical outcome measure by OMERACT (de Latour, Dalbeth & Taylor, 2015), as until recently there has been no widely agreed definition of gout flare used in the research literature and no consensus on how flares should be measured as an outcome (Gaffo et al 2012; de Latour, Dalbeth & Taylor, 2015; Taylor et al 2009). The recent definition of flares validated for use in clinical research (Gaffo et al 2018) was validated prior to the development of this study and requires data about clinical features not included in the questionnaire. As a definition of a flare was not included in the questionnaire this could have led to the potential for misclassification bias. Also, since the outcome measure used in this analysis was self-reported, and as participants were asked to recall the number of flares experienced in a reference period over 12- or 6- months, the potential for recall bias cannot be ruled out. However, patient-reported gout flare frequency,

as used in this study, has been identified as a feasible method in clinical trials (de Latour, Dalbeth & Taylor, 2015). Self-reported flares have also been compared to the criterion of an assessment by a rheumatologist and shown to have a sensitivity of 91% and negative predictive value of 96% (Gaffo et al 2012). There are several advantages to this self-reported approach to assess flares. Self-reports of flares can avoid the limitations of alternative methods of collecting gout flare data; such as using data derived from medical record review which often under-report the frequency of flares, as gout flares are often self-managed without seeking medical attention (Macfarlane et al 2016; Neogi et al 2006; Rothenbacher et al 2011; Sarawate et al 2006).

The mean number and variation in flares were calculated in this chapter based on the frequency of flares reported in categories. Thus, a limitation of these calculations is that the category ≥5 was taken to indicate five gout flares. Consequently, the mean number of flares, along with the variance, presented in this chapter is likely to be lower than if the flares were not reported within categories. However, as flares were reported in categories this justifies the decision made to model the gout flares data as ordinal data.

Latent growth curve modelling

An acknowledged limitation of latent growth curve modelling is that it follows the assumption that all individuals come from a single population and that a single latent growth curve can adequately describe growth parameters for the entire population (Jung & Wickrama, 2008; Ram & Grimm, 2009). The analysis undertaken in this chapter has identified a LGCM which describes the pattern of gout flares for 1164 participants, however the findings in this chapter (spaghetti plots, descriptive analysis of gout flares) also show that not all participants report similar patterns of flares over-time. Despite the limitations of LGCM it was important to determine the functional form of the latent growth curve for the gout flare data, in order to

inform the type of models used in the latent class growth analysis in the next chapter (Curran & Hussong, 2003; Curran, Obeidat & Losardo, 2010).

There are limitations to using multiple indices to assess model fit in latent growth curve modelling. Multiple indices were used in this chapter as the reporting of a variety of indices has been advocated due to the lack of consensus on the optimal model fit indices (Hooper, Coughlan & Mullen, 2008). A limitation of this approach is that there is no agreement about which of these multiple indices should carry more weight or be used in preference to other criteria. Using a range of criteria may lead to researcher bias if model fit indices are selected to support a particular viewpoint (Hooper, Coughlan & Mullen, 2008).

The latent growth curve modelling undertaken in this chapter took an exploratory approach to latent growth curve modelling; trialling various ways of modelling the data. The exploratory approach taken in this chapter is justified as there has not been previous research involving latent variable modelling of flare data. The findings from this chapter, relating to the quadratic ordinal model, informed the analysis in the next chapter of this thesis which aimed to identify classes (groups) of participants within this cohort who reported similar gout flare trajectories.

5.6 Conclusion

In conclusion, this chapter described self-reported gout flares over time within this prospective cohort. The frequency of self-reported flares, along with the diverse range in individual patterns of change in flares over time (trajectories), are presented. At each time-point, the largest proportion of participants reported experiencing no flares, however over a quarter of the cohort reported two or more flares. Various approaches to modelling the longitudinal flare data for the whole cohort were undertaken and more favourable results were returned for the quadratic ordinal latent growth curve model. This quadratic ordinal latent growth curve model was used in this chapter to describe the whole cohort. The findings of this chapter will be used to inform the analysis in the following chapter, where classes

(groups) of participants who reported similar flare trajectories will be identified within this cohort using latent class growth analysis.

6 Chapter Six Gout flare trajectory classes

6.1 Overview of chapter and aim

In the previous chapter the frequency of self-reported gout flares and the overall pattern of change (trajectory) of gout flares over three years for the whole cohort were presented. The aim of this chapter is to identify and describe latent classes (groups) of distinct gout flare trajectories in this prospective cohort of people living with gout in primary care. Thus, in the following chapter classes (groups) of people living with gout who have self-reported similar patterns of gout flares over time will be identified and the characteristics of class members described.

6.2 Background

The latent growth curve modelling undertaken in the previous chapter selected a quadratic ordinal model to describe the gout flare data for the whole cohort. However, the previous chapter also demonstrated that there was variation in the gout flare trajectories of individual participants within this cohort. Latent class growth analysis (LCGA) and growth mixture modelling (GMM) can enable classes (groups) of participants with similar trajectories to be identified (Muthen & Muthen, 2000). Once classes have been identified profiles describing the characteristics of each class can be created (Nagin, 1999).

6.3 Method and Analysis Plan

6.3.1 Aims and Objectives

Aim

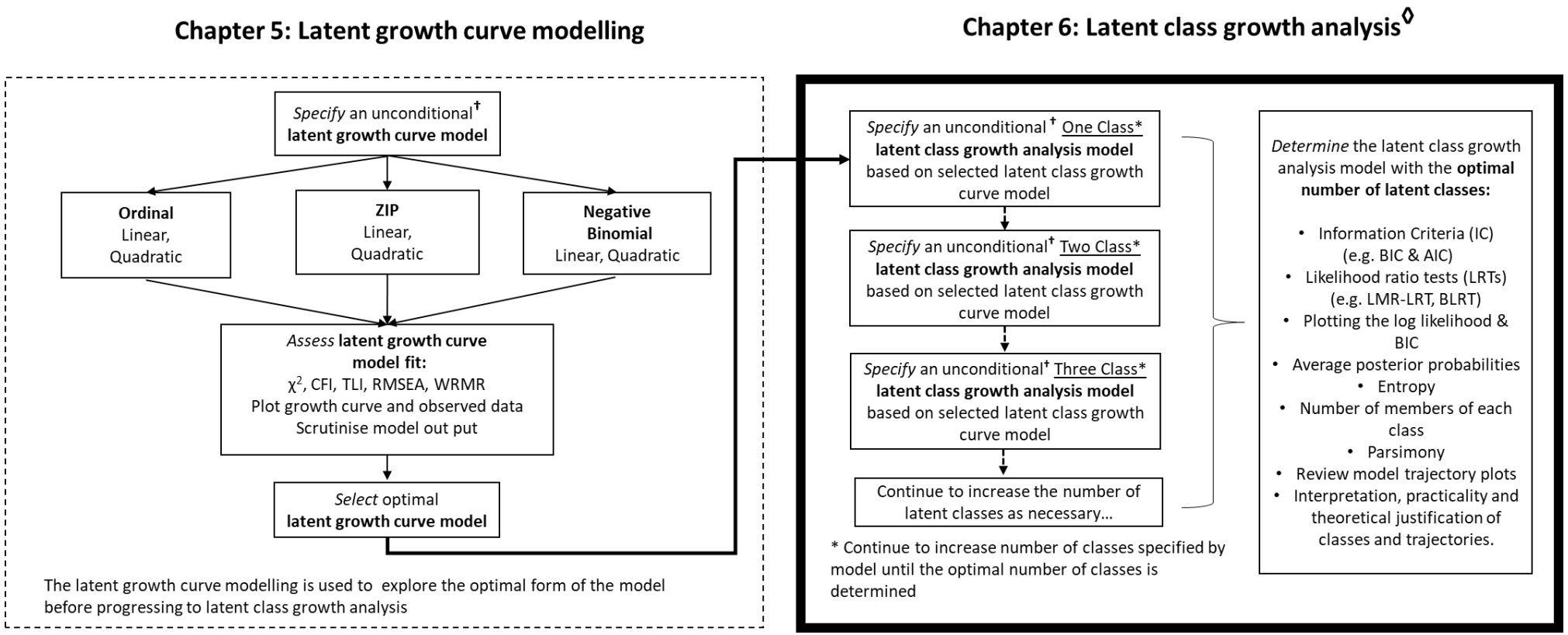
The aim of this chapter is to identify and describe latent classes (groups) of distinct gout flare trajectories in a prospective cohort of people living with gout in primary care.

Objectives

The objectives of this chapter are to:

1. Use latent class growth analysis (LCGA) and growth mixture modelling (GMM) to identify distinct latent classes of gout flare trajectories in people living with gout in primary care.
2. Describe the characteristics of members of each gout flare trajectory class.

A flowchart summary of the stages in the LCGA used in this chapter's methods, and how this follows on from the latent growth curve modelling in the previous chapter, can be found in Figure 6.1 on the next page.



◊ In latent class growth analysis the variance of intercepts and slopes are fixed to zero. Growth mixture modelling (where intercept and slope variance is permitted) was also undertaken following the same stages. †unconditional model without covariates

AIC Akaike Information Criteria; **BIC** Bayesian information criteria; **BLRT** Bootstrap likelihood ratio test; **CFI** Comparative Fit Index; **LMR-LRT** Lo Mendell Rubin likelihood ratio test; **TLI** Tucker-Lewis index; **RMSEA** Root mean square error of approximation; **WRMR** Weighted root-mean-square residual; **χ2** Chi-Squared Statistic; **ZIP** zero-inflated Poisson

Figure 6.1 Flowchart summary of latent class growth analysis and latent growth curve modelling

6.3.2 Data source

The data source for this analysis was derived from the cohort study described in chapter three. The primary observed variable used for the analysis in this chapter was the number of gout flares reported by participants at each questionnaire time-point (baseline, 6, 12, 24 and 36 months). Data related to gout-specific characteristics, medications, comorbidities, socio-demographics, BMI and alcohol consumption were taken from the questionnaire responder database, and medical record review database.

6.3.3 Sample size

No widely accepted sample size has been recommended for use in latent class growth analysis (Berlin, Parra & Williams, 2014). In simulation studies, the sample size requirements for latent growth curve modelling have been shown to be strongly influenced by the size of the regression coefficient, the degree of missing data and the inclusion of covariates (Muthen & Muthen, 2002). A simple growth curve with no covariates or missing data requires a sample of only 40 to obtain a sufficient power, whilst a conditional growth curve (with covariates) with missing data and a regression coefficient of 0.2 requires a sample of 250 to achieve a power of 0.8 (Muthen & Muthen, 2002). This variation in sample size requirement, dependent on the degree of missing data, has also been described by more recent simulation studies investigating optimal sample size in structural equation modelling (Wolf et al 2013).

Previous studies have successfully used patient reported outcomes to identify three to five latent class trajectories through LCGA in clinical cohorts in primary care with sample sizes of approximately 500 participants (Nicholls et al 2014; Rzewuska et al 2015). Thus, the analysis in this chapter, which is an unconditional analysis and has a sample size of 1164, is expected to be an adequate sample size to achieve appropriate power. However, due to the exploratory nature of the LCGA and GMM, the size of the regression coefficient required to reject the null hypothesis that the regression coefficient is zero was not known.

6.3.4 Missing data and sensitivity analysis

The extent of missing data in the gout flare data was presented in chapter four and chapter five. The LCGA and GMM in this chapter was undertaken using the statistical programme M*plus*. LCGA and GMM in *Mplus* handle missing data of the observed outcome variable by the use of full information maximum likelihood (FIML) estimation (Berlin, Parra & Williams, 2014; Jung & Wickrama, 2008; Muthen & Muthen, 2017; Ram & Grimm, 2009). FIML uses available observed outcome responses to supplement missing outcome responses to produce parameter estimates (Little et al 2014).

The LCGA was undertaken in the first instance by including all participants who self-reported the number of gout flares they had experienced on at least one time-point (n=1164). After this first LCGA, a sensitivity analysis was undertaken to see if including participants who reported flares at more time points yielded different trajectory patterns or optimal class number. Thus the LCGA was repeated with two other datasets: i) participants with observed outcome data (number of gout flares) at three or more time points (baseline plus two additional time points) (n=729) and ii) participants with complete data (an outcome reported at all five time points) (n=437). The findings from the three analyses were compared. After LCGA was undertaken, the extent of missing data within each class at each time point was described.

6.3.5 Prior to latent class growth analysis

The previous chapter (chapter five) details the descriptive analysis and latent growth curve modelling of gout flares which was undertaken in order to inform the analysis in this chapter.

6.3.6 Latent class growth analysis

In the absence of previously published latent variable modelling involving gout flares, LCGA was used in an exploratory manner in this chapter (Belin, Parra & Williams, 2014; Jung & Wickrama, 2008; Ram & Grimm, 2009). The optimal latent growth curve model functional

form determined in the previous chapter, the ordinal quadratic latent growth curve model, was used for the LCGA in this chapter (Berlin, Parra & Williams, 2014; Jung & Wickrama, 2008; Ram & Grimm, 2009; Wickrama et al 2016).

Specifying an unconditional latent class growth model with one-class only

LCGA utilises a forward classifying approach which commences with specifying a LCGA model which yields one latent class (Twisk & Hoekstra, 2012). Consequently, the initial stage in this latent class growth analysis was to specify an unconditional LCGA model with one class only, where within class variances are fixed to zero (Jung & Wickrama, 2008; Muthen & Muthen, 2000; Wickrama et al 2016) (see figure Figure 6.1).

After running this one-class LCGA model the output from the model was recorded and considered in line with the criteria described in Table 6.1.

Specifying unconditional latent class growth models with more latent classes

After the specification of a one-class LCGA model, the model was amended to estimate a two-class LCGA model. An iterative process then continued, sequentially increasing the number of classes within the LCGA model, until the optimal number of classes was achieved (see Table 6.1 and Figure 6.1) (Wickrama et al 2016). This continued estimation of as many classes as required until the optimal number of classes is achieved has been advocated as an exploratory approach in the investigation of LCGA (Berlin, Parra & Williams, 2014).

The starts and iterations were increased to the maximum, in order to be able to replicate the log likelihood (Muthen & Muthen, 2017).

After running each LCGA model, the output from the model was recorded and considered in line with the criteria described in the Table 6.1.

Determination of the optimal number of latent classes

A combination of different criteria need to be considered when determining the optimal number of latent classes in a LCGA analysis, due to the absence of one commonly accepted criterion (Nylund, Asparouhov & Muthen, 2007). The criteria considered whilst determining the optimal number of latent classes can be found in Table 6.1. The criteria used were selected because they have been recommended for use with LCGA and GMM, where the objective is to identify the model with the optimal number of classes and where the models have the same functional form (Berlin, Para & Williams, 2014; Jung & Wickrama, 2008). These criteria are different to the model fit criteria used in chapter five which were selected because they have been recommended in LGCM, where the objective is to evaluate the fit of models with different functional forms in order to decide which model is used to describe the whole cohort (Berlin, Para & Williams, 2014; Wu, West & Taylor, 2009).

Table 6.1 Guidance for the interpretation of criteria used to determine the optimal number of latent classes in LCGA

Criteria	Guidance
Bayesian information criteria (BIC)	A **lower BIC** indicated a better fitting model (Feldman, Masyn & Conger, 2009; Muthen & Muthen, 2000). The BIC has been advocated as the better performing information criteria when deciding the number of latent classes in LCGA in comparison to Akaike Information Criteria (AIC) (Nylund, Asparouhov & Muthen, 2007; Peugh & Fan, 2012).
Akaike information criteria (AIC)	A **lower AIC** result indicated a better fitting model (Nylund, Asparouhov & Muthen, 2007). The AIC has been found to be inconsistent and does not perform as well as the BIC when deciding the number of classes in LCGA (Nylund, Asparouhov & Muthen, 2007; Peugh & Fan 2012).
Lo Mendell Rubin Likelihood ratio Test (LMR-LRT)	**Statistically significant result** ($p < 0.05$) indicated an improvement in fit with the current model (*K*) versus a model with one less class (*K*-1) (Nylund, Asparouhov & Muthen, 2007; Peugh & Fan, 2012; Wickrama et al 2016). Due to inconsistencies with LMR-LRT (Jeffries, 2003), the LMR-LRT should not be used in isolation to decide the optimal number of latent classes (van de Schoot et al 2017).
Bootstrap Likelihood ratio test (BLRT)	**Statistically significant result** ($p < 0.05$) indicated an improvement in fit with the current model (*K*) versus a model with one less class (*K*-1) (Nylund, Asparouhov & Muthen, 2007; Peugh & Fan, 2012; Wickrama et al 2016). The **BLRT has been shown to outperform other tests** such as the BIC and LMR-LRT when determining the optimal number of latent classes (Nylund, Asparouhov & Muthen, 2007). The BLRT is associated with a significantly increased computational time and is more sensitive to non-normal distributions (Nylund, Asparouhov & Muthen, 2007; Wickrama et al 2016).
Plotting log likelihood and BIC	**Reviewing plots of the log likelihood and BIC** as the number of latent classes increased enables the point at which the log likelihood increases or BIC decreases to be identified on the graph (Muthen, 2001; van de Schoot et al 2017).
Average posterior probabilities	Average posterior probability results **nearer 1.0** indicate that individuals are assigned to a latent class with a high probability (Jung & Wickrama, 2008). Results greater than **0.70 to 0.80** indicate that the latent classes are grouping individuals with similar patterns of change, whilst separating individuals with different patterns of change (Andruff et al 2009).
Entropy	An entropy of 1.0 would indicate that each individual has a probability of membership of just one class (Feldman, Masyn & Conger, 2009). Entropy result cut offs of **1.0 (perfect), 0.8 (high), 0.60 (medium**), **0.4 (low)** have been proposed (Clark & Muthen, 2009). However, entropy was not originally intended to be used for model selection (Feldman, Masyn & Conger 2009) and high

	levels of entropy do not necessarily indicate that a model is clearly identifying trajectory classes (Clark & Muthen, 2009; Feldman, Masyn & Conger, 2009).
Number of members per class	A minimum number of members per latent class of less than 1% of the total cohort is insufficient and the number should ideally be >5% (Jung & Wickrama, 2008; Wickrama et al 2016).
Parsimony	Where there is uncertainty regarding the optimal number of latent classes based on the results from information criteria and likelihood ratio tests, then a model with less estimated parameters and less classes would be preferable (Feldman, Masyn & Conger, 2009).
Plot latent class growth analysis models	Plotting the growth trajectories of the classes in each model aided interpretability and identified very similar growth trajectories (Feldman, Masyn & Conger, 2009; Ram & Grimm, 2009).
Interpretation, practicality and theoretical justification of classes and trajectories.	The results of the LCGA must be interpreted within the context of the specific research question (Bauer & Curran, 2003; Feldman, Masyn & Conger 2009; Jung & Wickrama, 2008). The different potential latent growth trajectories were reviewed within the context of whether they can be practically interpreted as being realistic distinct gout flare trajectories.

Description of gout flare latent class growth analysis trajectories

After deciding on the optimal number of latent classes of flare trajectories, the final gout flare latent class trajectories were presented in graphical form. As the data were ordinal, the probability of reporting two or more flares at each time point was plotted rather than the mean number of flares. Each latent class trajectory was allocated a descriptive title which described the pattern in flares over time displayed by the latent class growth trajectory. These descriptive names were derived following discussion with the supervisory team which included clinicians and statisticians. The trajectory classes were also presented to and discussed with both a wider rheumatological and general practice audiences at regional meetings and national conferences.

Consultation with patient and public involvement

The proposed latent trajectories were presented and discussed with a patient and public involvement and engagement group who were lay members of the Keele University Research User Group (RUG) living with gout. The group included both female and male members who had been living with gout for different periods of time. Individuals shared their views on the descriptive names assigned to the latent classes, along with the interpretability of the classes and class characteristics.

Characteristics and profile of members of different trajectory classes

After the identification of distinct latent class trajectories, descriptive statistics for each of the different latent classes, relating to baseline gout-specific variables, medications, comorbidities, socio-demographic variables, BMI and alcohol consumption, were derived. Nominal and ordinal categorical baseline variables were analysed by undertaking frequency counts and calculating percentages. Continuous numerical variables were analysed by

calculating the mean and standard deviation. These descriptive characteristics of trajectory class members (profiles) for different classes were then compared (Nagin, 1999).

6.3.7 Growth mixture modelling

A quadratic ordinal Growth Mixture Model (GMM) was explored. Unlike the LCGA model, in a GMM typically the latent intercept, linear slope and quadratic slope variances are not all fixed to zero and thus covariance is permitted (Berlin, Parra & Williams, 2014; Jung & Wickrama, 2008; Wickrama et al 2016). An iterative process of increasing the number of classes in the GMM, using the criteria in Table 6.1 to assess the optimal number of classes, was followed.

6.4 Results

6.4.1 Latent class growth analysis

1164 participants had reported the number of gout flares they had experienced on at least one time point. Table 6.2 displays the model fit indices for the different LCGA models with increasing number of classes. The lowest BIC was returned for the six-class LCGA model, the lowest AIC was returned for the eight-class model. However, the four, seven, and eight-class models displayed warnings as indicated in Table 6.2. A statistically significant BLRT result was returned for all LCGA models with one to eight classes. Statistically significant LMR-LRT results were achieved for models of five classes or less. The highest entropy score, 0.703, was returned for the two-class model with 0.649 and 0.648 returned for the three-class model and six-class models respectively. In the plot of the change in log likelihood (Figure 6.2) and BIC (Figure 6.3) as the number of classes increased, the log likelihood continued to rise, whereas the lowest BIC was seen in the LCGA model with six classes. Figure 6.4 to Figure 6.11 display plots of the one, two, three, four, five, six, seven, and eight-class LCGA models respectively.

The six-class model was selected as the optimal LCGA model. This model returned the lowest BIC and a statistically significant BLRT. When selecting the optimal number of latent classes

the BIC has been advocated as the better performing information criterion and the BLRT has been shown to out-perform other tests (Nylund, Asparouhov & Muthen, 2007). The six-class model also returned medium to high entropy, which was higher than all models except the two and three-class models. The six-class model yielded trajectory patterns which could reflect the range of potential gout flare trajectories in clinical practice. Whilst the two and three class models returned higher posterior probabilities and entropy results, these models yielded gout flare trajectory patterns which displayed little change over time, therefore not fully reflecting the range of patterns of change in flares identified in the previous chapter. The two and three-class models yielded less favourable BIC results.

Table 6.2 Model fit results for LCGA (participants who reported gout flares at ≥1 time-points)

Number of classes	AIC	BIC	BLRT	LMR-LRT	Entropy	Average posterior probabilities	Number per class (%) based on most likely class membership
1	11200.304	11235.722	-	-	-		1164 (100)
2	10096.374	10152.029	P value <0.001	P value <0.001	0.703	0.884/0.933	460/704 (39.5/60.5)
3	9900.376	9967.270	P value <0.001	P value <0.01	0.649	0.794/ 0.865/0.808	369/603/192 (31.7/51.8/16.5)
4[a]	9843.368	9939.501	P value <0.001	P value <0.01	0.623	0.790/0.757/0.758/0.694	360/299/409/96 (30.9/25.7/35.1/8.3)
5	9799.746	9916.117	P value <0.001	P value <0.05	0.616	0.744/0.761/0.724/0.769/0.665	354/140/330/251/89 (30.4/12.0/28.4/21.6/7.6)
6	9765.561	**9902.171**	**P value <0.001**	P value 0.1110	0.648	0.750/0.709/0.644/0.768/0.815/0.758	349/276/95/287/14/143 (29.9/23.7/8.2/24.7/1.2/12.3)
7[a,b]	9757.126	9913.974	P value <0.001 [b]	P value 0.0967	0.635	0.652/0.683/0.706/0.788/0.743/0.733/0.734	95/37/280/12/125/350/265 (8.2/3.1/24.1/1.0/10.7/30.1/22.8)
8[a,b,c]	9746.744	9923.830	P value <0.001[b,c]	P value 0.1121	0.602	0.698/0.663/0.713/0.726/0.581/0.687/0.752/ 0.552	37/118/344/125/226/223/13/78 (3.2/10.1/29.6/10.7/19.4/19.1/1.1/6.7)

AIC, Akaike Information Criteria; **BIC**, Bayesian Information Criteria; **BLRT**, Bootstrap likelihood ratio test; **LMR-LRT**, Lo Mendell Rubin likelihood ratio test.
The **lowest BIC** and **statistically significant BLRT** for the six-class solution indicated the optimal number of classes.
[a] model fixed to avoid singularity, warning given that model may not be identified, [b] Warning that not all of bootstraps converged, [C] Warning that some draws had smaller LRT than observed LRT
Table formatted as per Guidelines for Reporting on Latent Trajectory Studies (GRoLTs) (van de Schoot et al 2017)

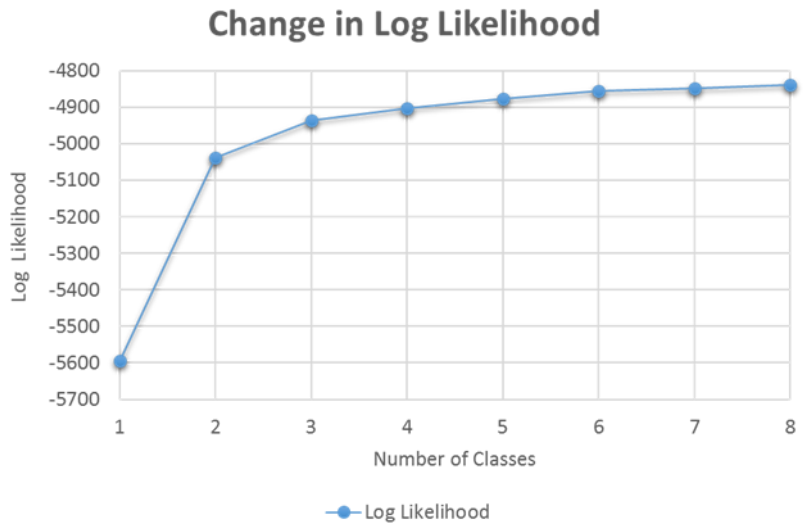


Figure 6.2 Plot of change in log likelihood for LCGA models with different number of classes

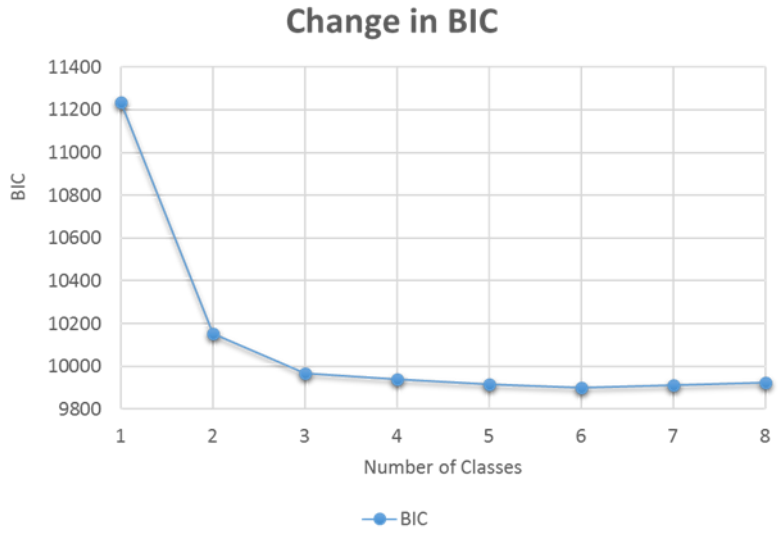


Figure 6.3 Plot of change in BIC for LCGA models with different number of classes

Figure 6.4 to Figure 6.11 display the plots of the one, two, three, four, five, six, seven, and eight-class LCGA models respectively. In each plot, the probability of two or more gout flares is plotted on the Y axis and the growth curve of classes for the estimated model is shown. The plot of LCGA models with five and six classes both show class trajectory plots where the probability decreases steeply and the six class LCGA model also displays a trajectory class where the probability is shown to slowly increase over time.

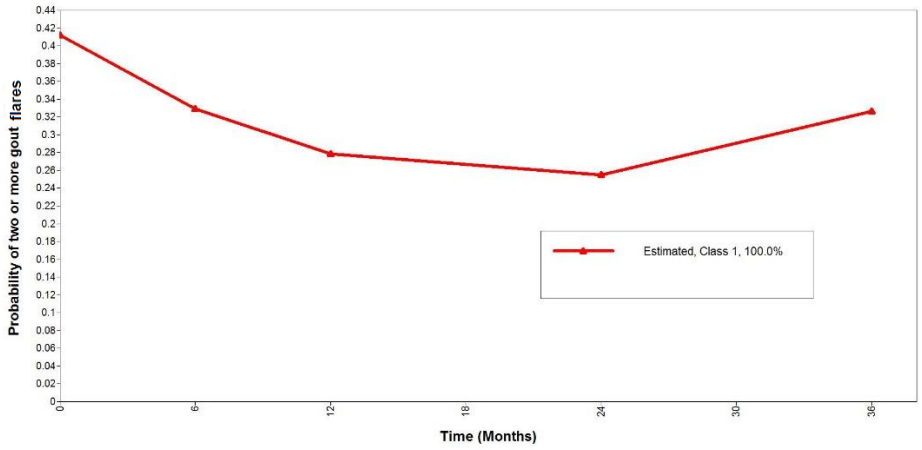


Figure 6.4 One class quadratic ordinal LCGA model plot; probability of ≥2 gout flares at each time-point

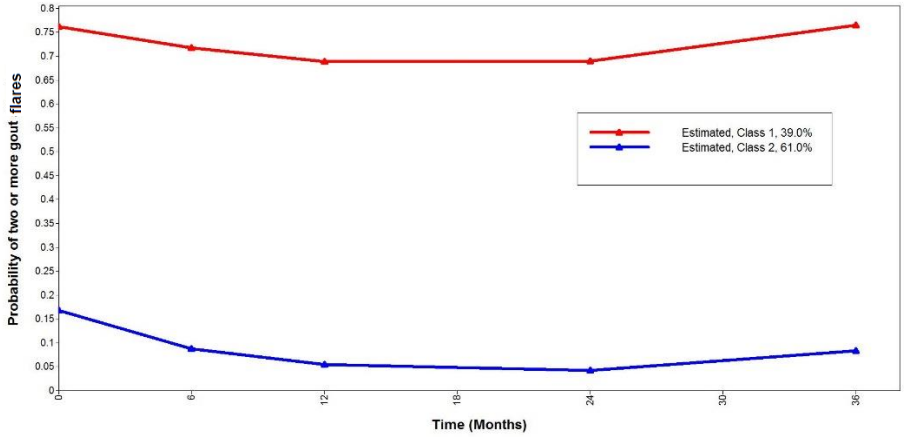


Figure 6.5 Two class quadratic ordinal LCGA model plot; probability of ≥2 gout flares at each time-point

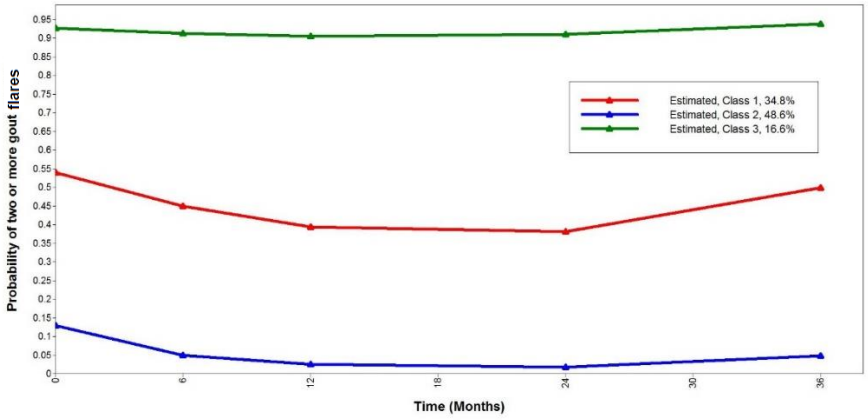


Figure 6.6 Three class quadratic ordinal LCGA model plot; probability of ≥2 gout flares at each time-point

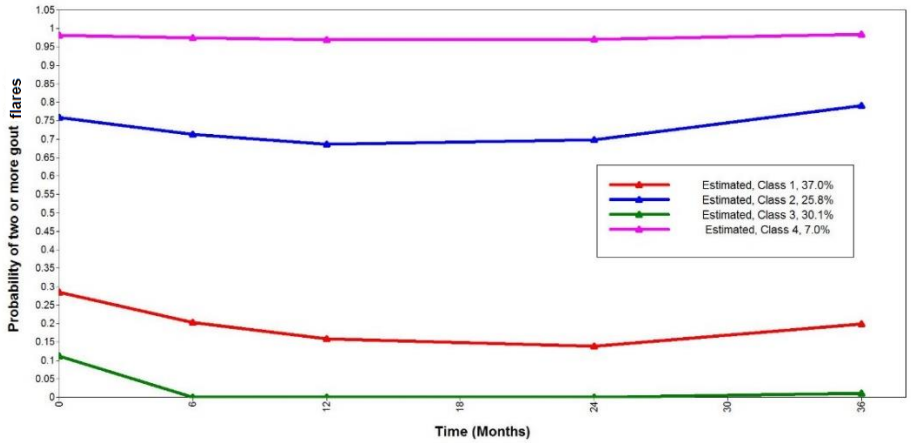


Figure 6.7 Four class quadratic ordinal LCGA model plot; probability of ≥2 gout flares at each time-point

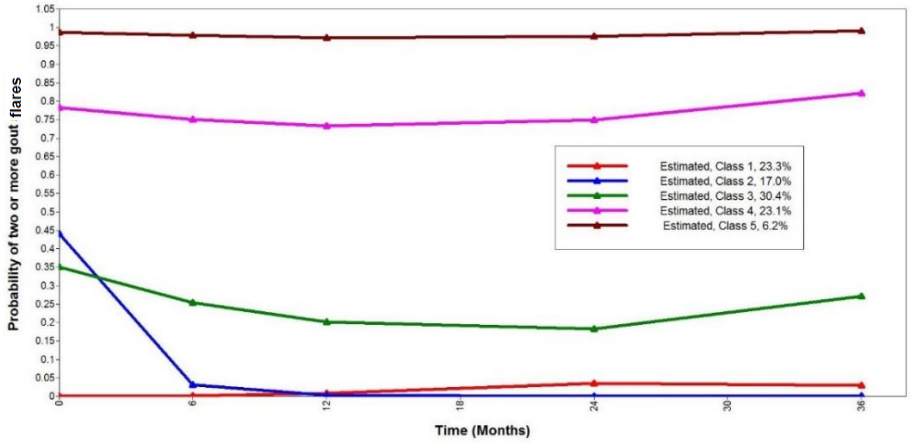


Figure 6.8 Five class quadratic ordinal LCGA model plot; probability of ≥2 gout flares at each time-point

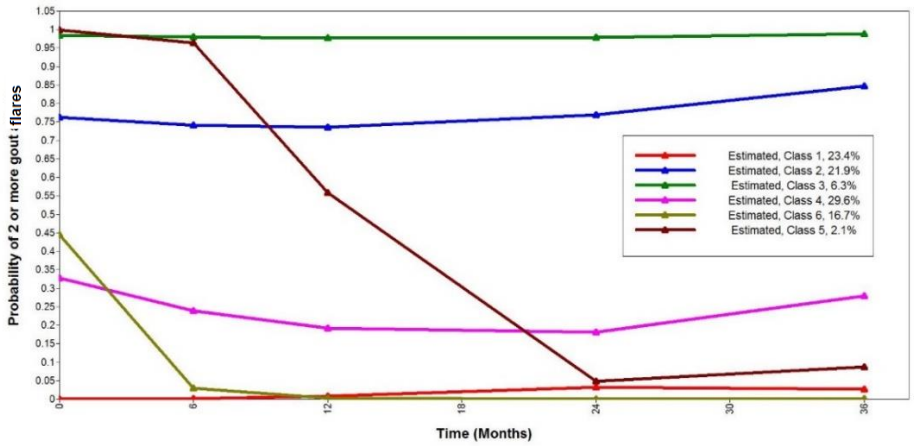


Figure 6.9 Six class quadratic ordinal LCGA model plot; probability of ≥2 gout flares at each time-point

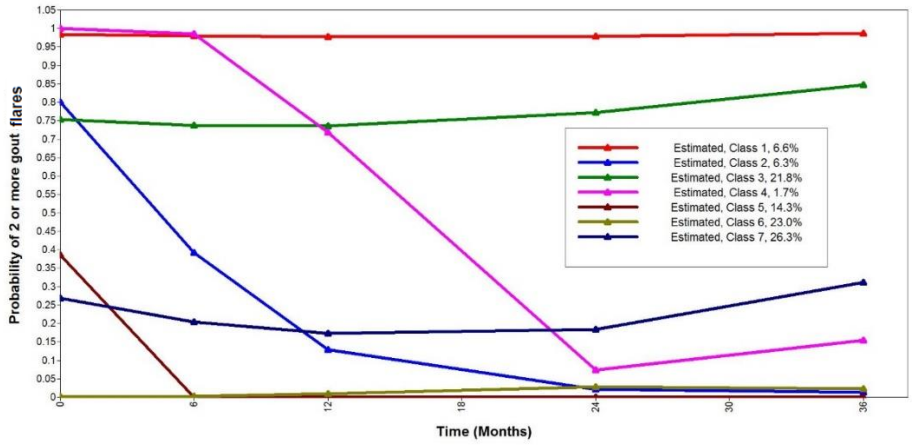


Figure 6.10 Seven class quadratic ordinal LCGA model plot; probability of ≥2 gout flares at each time-point

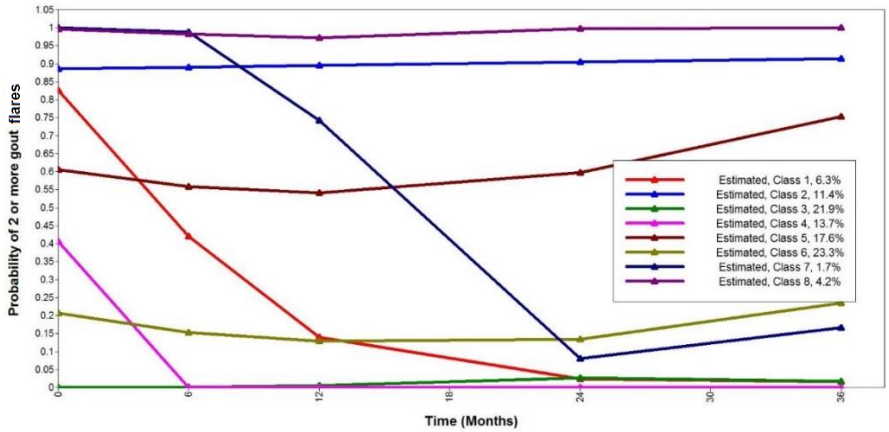


Figure 6.11 Eight class quadratic ordinal LCGA model plot; probability of ≥2 gout flares at each time-point

6.4.2 Descriptive names for latent trajectory classes

Figure 6.12 displays a plot of the estimated model class trajectories for the six class LCGA and the descriptive name allocated to each class. The classes were 'Frequent and Persistent' n=95, 'Gradually worsening' n=276, 'Frequent then improving' n=14, 'Moderately frequent' n=287, 'Moderately frequent then improving' n=143, and 'Infrequent' n=349. Individual spaghetti plots were also created (see appendix 11) displaying the gout flares reported for each participant in each class and an interpolation line fitted, the pattern of these spaghetti plots were also considered when creating the descriptive names for the classes.

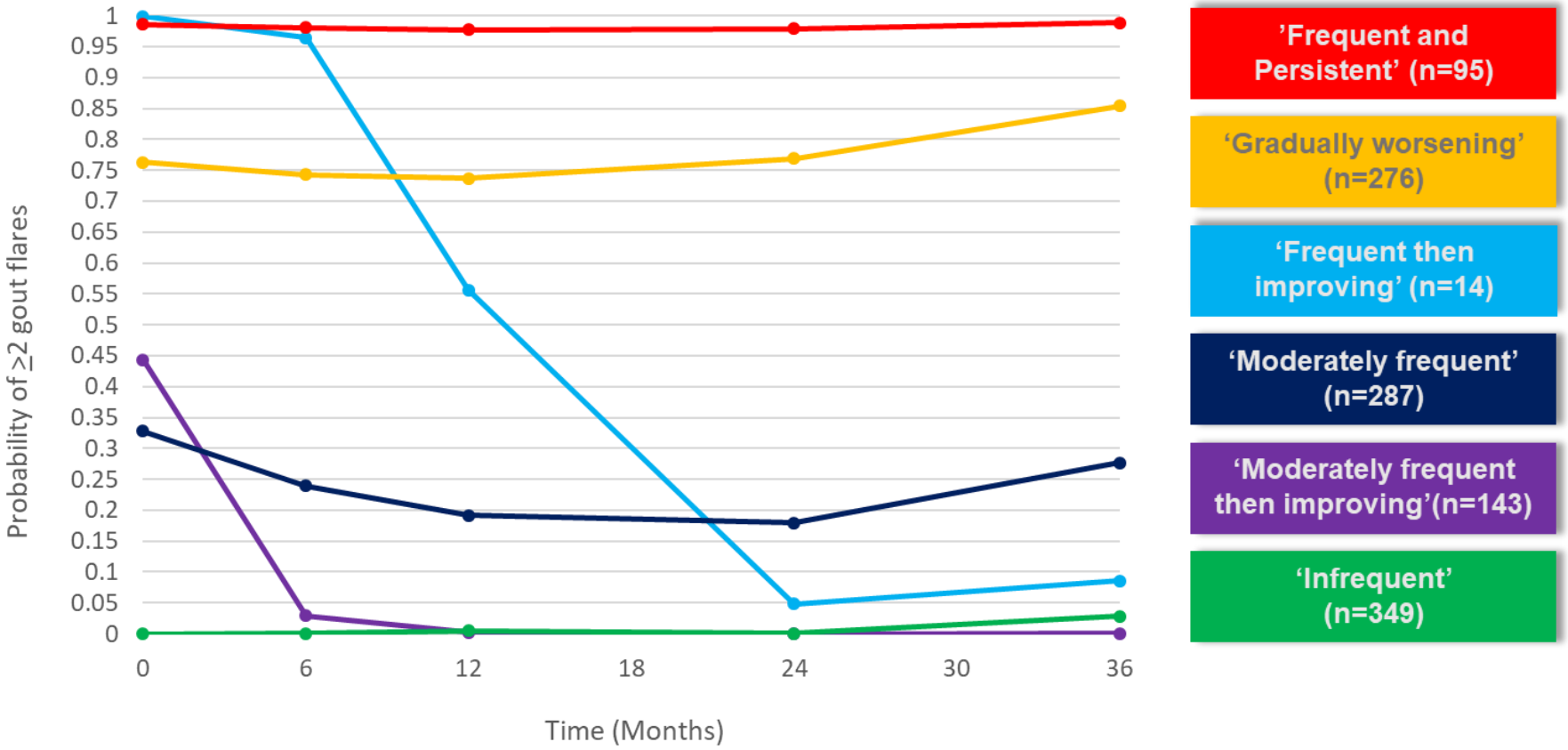


Figure 6.12 Plot of six class ordinal quadratic LCGA model with descriptive names and class sizes

6.4.3 Missing data and sensitivity analysis

Missing data due to unit and item non-response per class

The percentage of missing gout flare data due to unit non-response (where the questionnaire was not returned) in all classes was 0% at baseline but ranged from 21.4% to 69.5% at 36 months (see Table 6.3). From six months onwards the 'frequent then improving' had the lowest proportion of missing data at all time-points compared to other classes, whereas the 'frequent and persistent' had the highest proportion of missing data.

Table 6.3 Missing data due to unit non-response per class for six-class LGCA model

Missing data per class self-reported gout flares	**'frequent and persistent'** n=95	**'gradually worsening'** n=276	**'frequent then improving'** n=14	**'moderately frequent'** n=287	**'moderately frequent then improving'** n=143	**'infrequent'** n=349
Baseline n(%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
6 months n(%)	45 (47.4)	109 (41.3)	1 (7.1)	70 (24.4)	11 (7.7)	110 (31.5)
12 months n(%)	53 (55.8)	124 (44.9)	1 (7.1)	85 (29.6)	37 (25.9)	143 (41.0)
24 months n(%)	56 (58.9)	126 (45.7)	1 (7.1)	93 (32.4)	47 (32.9)	145 (41.5)
36 months n(%)	66 (69.5)	144 (52.2)	3 (21.4)	120 (41.8)	62 (43.1)	164 (47.0)

The percentage of missing gout flare data due to item non-response (where there was no response to the gout flare item in a returned questionnaire) in the different classes ranges from 0 to 7.1% at baseline and from 0 to 2.9% at 36 months (Table 6.4). The proportion of missing data due to item non-response was far less than the proportion of missing data due to unit non-response (Table 6.3).

Table 6.4 Missing data due to item non-response per class for six-class LGCA model

Missing data per class self-reported gout flares	**'frequent and persistent'** n=95	**'gradually worsening'** n=276	**'frequent then improving'** n=14	**'moderately frequent'** n=287	**'moderately frequent then improving'** n=143	**'infrequent'** n=349
Baseline n(%)	3 (3.2)	10 (3.6)	1 (7.1)	11 (3.8)	0 (0.0)	16 (4.6)
6 months n(%)	3 (3.2)	5 (1.8)	0 (0.0)	6 (2.1)	1 (0.7)	2 (0.6)
12 months n(%)	1 (1.1)	4 (1.4)	0 (0.0)	5 (1.7)	4 (2.8)	6 (1.7)
24 months n(%)	2 (2.2)	7 (2.5)	1 (7.1)	8 (2.8)	4 (2.8)	9 (2.6)
36 months n(%)	0 (0.0)	8 (2.9)	0 (0.0)	7 (2.4)	3 (2.1)	7 (2.0)

Sensitivity analysis

For the different LCGA models using the dataset consisting of participants who self-reported gout flares at three or more time-points (n=729), the lowest BIC was returned for the six-class LCGA model, the lowest AIC was returned for the eight-class model (Table 6.5). A statistically significant BLRT result was returned for all LCGA models with one to eight classes. Statistically significant LMR-LRT results were achieved for models of four classes or less and also the eight-class model (although the eight-class model displayed a warning about the LRT result). The highest entropy score, 0.850, was returned for the two-class model with 0.783 and 0.777 returned for the three-class model and six-class models respectively.

In the plot of the change in log likelihood and BIC in participants who self-reported gout flares at three or more time-points, as the number of classes increased the log likelihood continued to rise, whereas the lowest BIC was seen in the LCGA model with six classes. (Figure 6.13 and Figure 6.14 respectively).

The plot of the six-class LCGA model in this dataset (Figure 6.15) displayed the same trajectories observed in the LCGA including all participants who self-reported gout flares on at

least one time-point. The plots of all the LCGA models (one to eight classes) for this dataset can be found in appendix 12.

The LCGA including participants who reported self-reported gout flares at three or more time-points also identified the six-class model as the optimal LCGA model; as the six-class model yielded the lowest BIC, a statistically significant BLRT and clinically interpretable trajectory patterns.

Table 6.5 Model fit results for LCGA (participants who reported gout flares at ≥3 time-points)

Number of classes	AIC	BIC	BLRT	LMR-LRT	Entropy	Average posterior probabilities	Number per class (%) based on most likely class membership
1	9202.198	9234.113	-	-	-	-	729 (100)
2	8155.788	8206.297	P value <0.001	P value <0.001	0.850	0.969/0.942	457/272 (62.7/37.3)
3	7971.349	8040.224	P value <0.001	P value <0.05	0.783	0.859/0.894/0.931	239/114/376 (32.8/15.6/51.6)
4[a]	7919.649	8006.891	P value <0.001	P value <0.05	0.763	0.862/0.872/0.847/0.879	46/174/255/254 (6.3/23.8/35/34.9)
5	7877.637	7983.245	P value <0.001	P value 1 .000	0.749	0.845/0.909/0.861/0.823/0.841	112/32/164/215/206 (15.4/4.4/22.5/29.5/28.2)
6	7843.885	**7967.860**	**P value <0.001**	P value 0.0940	0.777	0.877/0.829/0.895/0.844/0.862/0.853	34/213/11/208/153/110 (4.7/29.2/1.5/28.5/21/15.1)
7[b,c]	7834.252	7976.594	P value <0.001	P value 0.3076	0.726	0.710/0.834/0.820/0.750/0.780/0.831/00.842	115/32/105/80/178/206/13 (15.8/4.4/14.4/10.9/24.4/28.3/1.8)
8[b,c]	7821.807	7982.516	P value <0.001	P value 0.0338	0.737	0.738/0.817/0.831/0.777/0.829/0.760/0.83/ 0.712	22/105/3/168/205/59/14/125 (3/14.4/4.3/23.1/28.1/8.1/1.9/17.1)

AIC, Akaike Information Criteria; **BIC**, Bayesian Information Criteria; **BLRT**, Bootstrap likelihood ratio test; **LMR-LRT**, Lo Mendell Rubin.
The **lowest BIC** and **statistically significant BLRT** for the six-class solution indicated the optimal number of classes
[a] model fixed to avoid singularity, warning given that model may not be identified, [b] Warning that not all of bootstraps converged,
[C] Warning that some draws had smaller LRT than observed LRT
Table formatted as per Guidelines for Reporting on Latent Trajectory Studies (GRoLTs) (van de Schoot et al 2017)

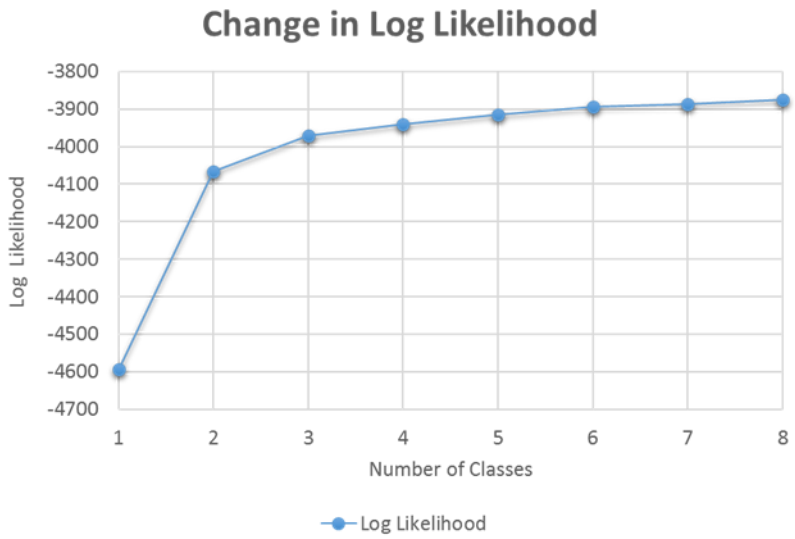


Figure 6.13 Plot of change in log likelihood for LCGA models with different number of classes (for participants who responded at ≥3 time-points)

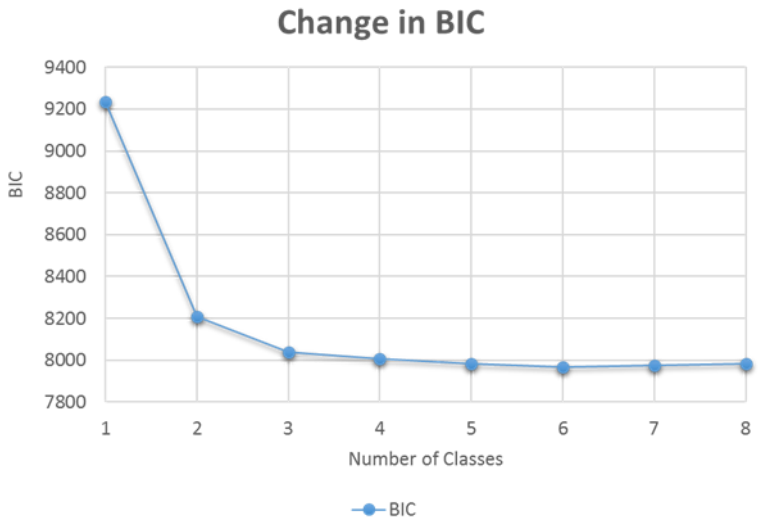


Figure 6.14 Plot of change in BIC for LCGA models with different number of classes (for participants who responded at ≥3 time-points)

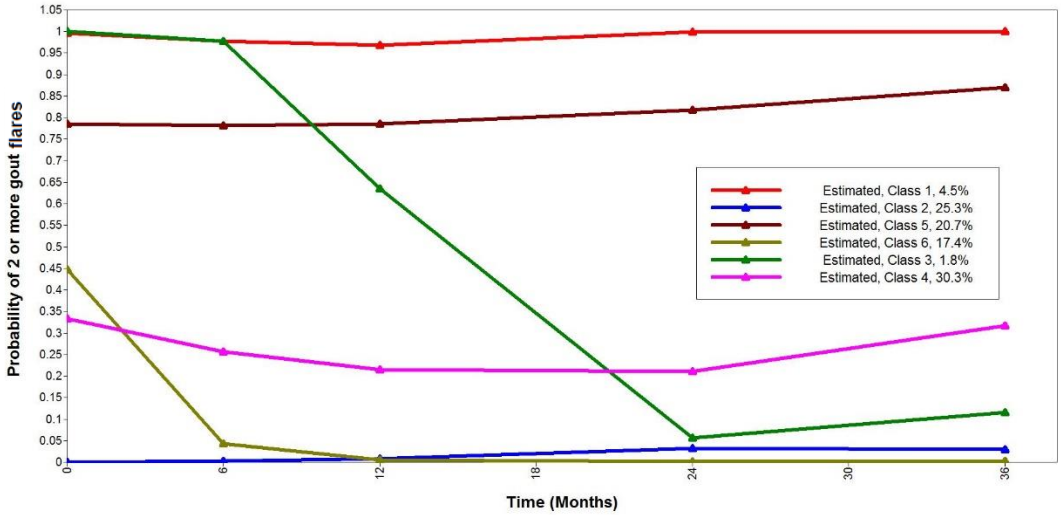


Figure 6.15 Six class quadratic ordinal LCGA model plot; probability of ≥2 gout flares at each time-point (for participants who responded at ≥3 time-points)

For the LCGA models in the dataset consisting of participants who self-reported gout flares at five time-points (n=437), the lowest BIC was returned for the six-class LCGA model, the lowest AIC was returned for the eight-class model (Table 6.6). A statistically significant BLRT result was returned for all LCGA models with one to eight classes. Statistically significant LMR-LRT results were achieved for the one-class, two-class and six-class LCGA model. The highest entropy score, 0.876 was returned for the two-class model. There were warnings for the seven and eight-class models that related to non-convergence of bootstraps and problems with the LRT results. In the plots of the change in log likelihood and BIC for the LGCA models (Figure 6.16 and Figure 6.17 respectively) the log likelihood continued to rise as the number of classes increased, whereas the lowest BIC was seen in the LCGA model with six classes.

The LCGA in participants who self-reported gout flares at five time-points also identified the six-class model as the optimal LCGA model; as the six-class model yielded the lowest BIC, a statistically significant BLRT and clinically interpretable trajectory patterns.

The plot of the six-class LCGA model in this dataset (Figure 6.18) displayed the same trajectories observed in the LCGA including all participants who self-reported gout flares on at least one time-point and also the dataset including all participants who self-reported gout flares on at least three time-points. The plots of all the LCGA models (one to eight classes) for this dataset can be found in appendix 12. This sensitivity analysis revealed that each LCGA in the three datasets displayed the same gout flare trajectory patterns for the six-class model.

Table 6.6 Model fit results for LCGA (participants who reported gout flares at all 5 time-points)

Number of classes	AIC	BIC	BLRT	LMR-LRT	Entropy	Average posterior probabilities	Number per class (%) based on most likely class membership
1	6017.515	6046.074	-	-	-	-	437 (100)
2	5262.490	5307.369	P value <0.001	P value <0.001	0.876	0.967/0.962	284/153 (65/35)
3	5140.591	5201.790	P value <0.001	P value <0.01	0.823	0.897/0.947/0.889	61/238/138 (14/54.5/31.5)
4	5102.776	5180.295	P value <0.001	P value 0.0509	0.872	0.901/0.955/0.908/0.892	129/238/59/11 (29.5/54.5/13.5/2.5)
5[a]	5067.561	5161.399	P value <0.001	P value 0.2814	0.821	0.907/0.931/0.894/0.857/0.892	141/10/29/95/162 (32.3/2.3/6.6/21.7/37.1)
6	5036.774	**5146.932**	**P value <0.001**	**P value <0.001**	0.799	0.853 /0.870/0.937/0.901/0.861/0.892	139/113/10/24/84/67 (31.8/25.9/2.3/5.5/19.2/15.3)
7[b,c]	5025.453	5151.931	P value <0.001	P value 0.3300	0.799	0.824/0.928/0.888/ 0.788/ 0.868/0.912/0.854	74/10/67/21/111/15/139 (16.9/2.3/15.3/4.8/25.4/3.4/31.8)
8[a,b,c]	5024.706	5167.503	P value <0.001	P value 0.1533	0.776	0.849/0.926/0.779/0.826/0.864/0.776/0.821/ 0.941	105/10/26/21/52/74/135/14 (24.0/2.3/6.0/4.8/11.9/16.9/30.9/3.2)

AIC, Akaike Information Criteria; **BIC**, Bayesian Information Criteria; **BLRT**, Bootstrap likelihood ratio test; **LMR-LRT**, Lo Mendell Rubin.
The **lowest BIC** and **statistically significant BLRT** for the six-class solution indicated the optimal number of classes
[a] model fixed to avoid singularity, warning given that model may not be identified, [b] Warning that not all of bootstraps converged,
[C] Warning that some draws had smaller LRT than observed LRT
Table formatted as per Guidelines for Reporting on Latent Trajectory Studies (GRoLTs) (van de Schoot et al 2017)

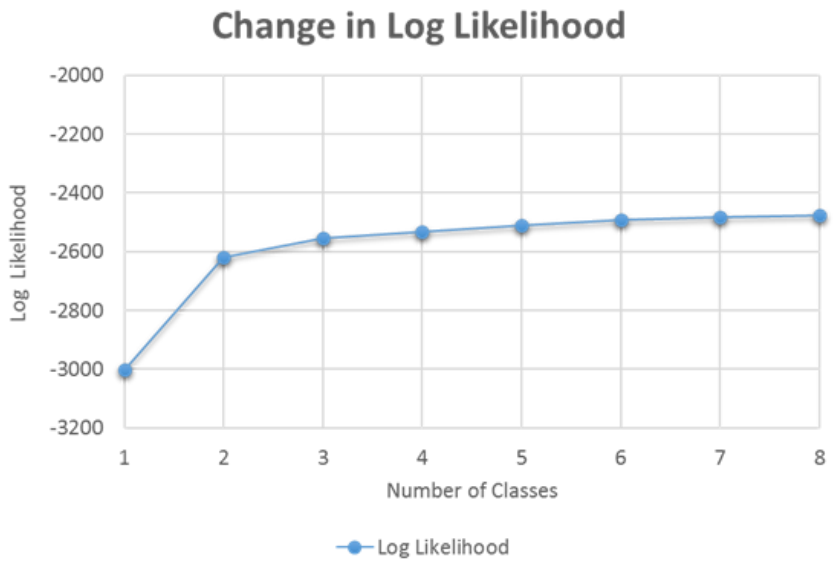


Figure 6.16 Plot of change in log likelihood for LCGA models with different number of classes (for participants who responded at 5 time-points)

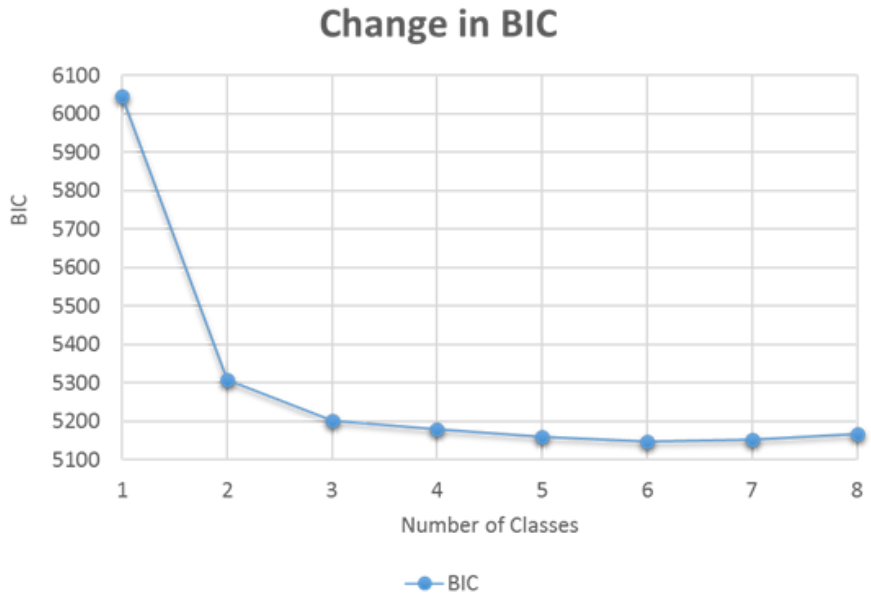


Figure 6.17 Plot of change in BIC for LCGA models with different number of classes (for participants who responded at 5 time-points)

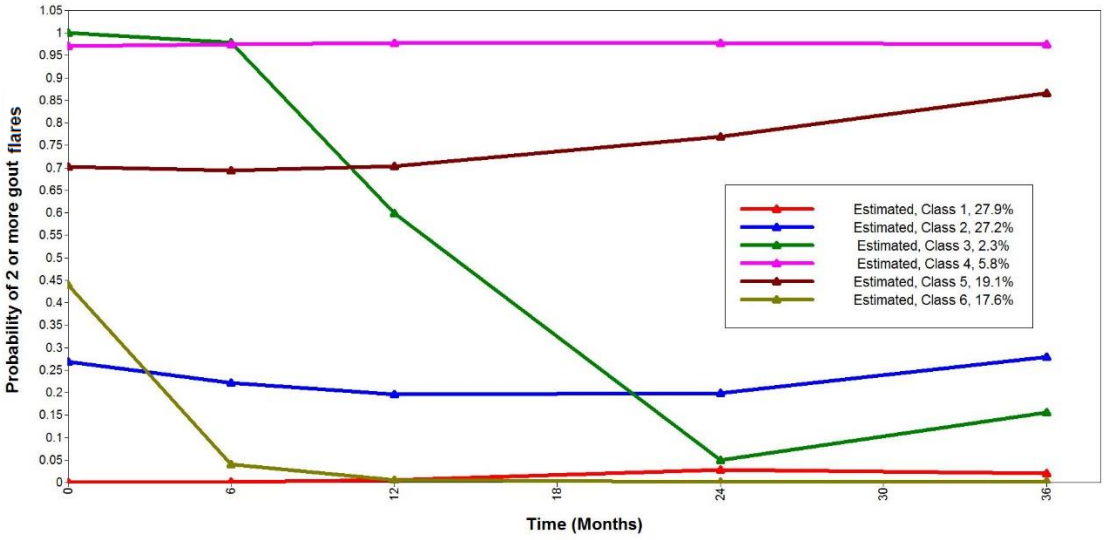


Figure 6.18 Six class quadratic ordinal LCGA model plot; probability of ≥2 gout flares at each time-point (for participants who responded at 5 time-points)

6.4.4 Descriptive characteristics of trajectory class members

Gout-specific characteristics

In the 'frequent and persistent', 'frequent then improving' and 'gradually worsening' classes 95.7%, 92.9% and 86.2% of class members respectively had experienced two or more gout flares in the 12 months prior to baseline, in contrast to 0% in the 'infrequent' class and 26.5% in the 'moderately frequent' class (Table 6.7). The 'infrequent' class had the longest mean (SD) disease duration of 15.6 (12.3) years, whilst the disease duration in the other classes ranged from 7.5 to 11.6 years. A larger proportion of the 'frequent and persistent' and 'frequent then improving' classes (68.4% and 78.6% respectively) had a history of self-reported oligo/polyarticular flares, compared with the 'infrequent', 'moderately frequent' and 'moderately frequent then improving' classes (26.6%, 30.3%, and 30.8% respectively). In addition, 4.2% and 7.1% of class members in the 'frequent and persistent' and 'frequent then improving' had tophi recorded in their medical records in the two years prior to baseline, compared with only 1.4% and 1.1% in the 'moderately frequent' and 'infrequent' classes respectively. The 'gradually worsening' class had the highest mean maximum serum urate

level (481 µmol/L) recorded in their medical records in the two years prior to baseline, followed by the 'frequent and persistent' class (469 µmol/L), whilst the 'infrequent' class had the lowest mean maximum serum urate level (377 µmol/L). The 'frequent and persistent' class had the highest proportion of class members (44.2%) with a maximum serum urate level greater than 360 µmol/L, in contrast with 15.2% in the 'infrequent' class.

Table 6.7 Gout specific characteristics of class members at baseline

Baseline	**All in LCGA** n=1164	**'frequent and persistent'** n=95	**'gradually worsening'** n=276	**'frequent then improving'** n=14	**'moderately frequent'** n=287	**'moderately frequent then improving'** n=143	**'infrequent'** n=349
0 gout flares†	398(34.2)	0 (0.0)	7(2.5)	0 (0.0)	55(19.2)	3(2.1)	333(95.4)
1 gout flare†	231(19.8)	1(1.1)	21(7.6)	0 (0.0)	145(50.5)	64(44.8)	0 (0.0)
2 gout flares†	187(16.1)	1(1.1)	93(33.7)	0 (0.0)	49(17.1)	44(30.8)	0 (0.0)
3 gout flares†	103(8.8)	2 (2.1)	72(26.1)	0 (0.0)	13(4.5)	16(11.2)	0 (0.0)
4 gout flares†	67(5.8)	4 (4.2)	46(16.7)	0 (0.0)	10(3.5)	7(4.9)	0 (0.0)
5 or more gout flares†	137(11.8)	84 (88.4)	27(9.8)	13 (92.9)	4(1.4)	9(6.3)	0 (0.0)
≥2 gout flares†	494(42.4)	91(95.7)	238(86.2)	13 (92.9)	76(26.5)	76(53.1)	0(0.0)
Missing gout flare	41(3.5)	3(3.2)	10(3.6)	1(7.1)	11(3.8)	0(0)	16(4.6)
Disease duration mean (SD)	11.9(12.0)	11.6(12.3)	11.0(11.6)	7.5(5.9)	9.6(11.2)	9.8(12.3)	15.6(12.3)
Occurrence of current flare	131(11.3)	48(50.5)	43(15.6)	8(57.1)	19(6.6)	10(7)	3(0.9)
History of oligo/polyarticular flares	432(37.1)	65(68.4)	132(47.8)	11(78.6)	87(30.3)	44(30.8)	93(26.6)
Record of tophi ◇	25(2.1)	4(4.2)	8(2.9)	1(7.1)	4(1.4)	4(2.8)	4(1.1)
Serum urate level >360 µmol/L ◇*	349(30.0)	42(44.2)	87(31.5)	6(42.9)	103(35.9)	58 (40.6)	53(15.2)
Maximum serum urate µmol/L mean (SD) ◇*	440.98 (115.2)	469.08 (116.6)	480.51 (113.3)	461.43 (185.5)	450.71 (102.8)	434.48 (105.7)	377.15 (107.7)
Missing Serum urate◇	707(60.7)	44(46.3)	174(63.0)	7(50.0)	164(57.1)	68(47.6)	250(71.6)

Values are n(%) unless stated otherwise

† **in previous 12 months** at baseline; ◇**In medical record** in the two years prior to baseline; ***highest serum urate** recorded

Medications

In the 'infrequent' class, 73% of class members reported taking allopurinol at baseline, in contrast with 28.6% in the 'frequent then improving' class (Table 6.8), whilst 48.4%, 46.7%, 40.1% and 52.4% of class members in the 'frequent and persistent', 'gradually worsening', 'moderately frequent' and 'moderately frequent then improving' respectively reported taking allopurinol. The 'infrequent' class also had the highest proportion of class members with a prescription for allopurinol in their medical records in the two years prior to baseline (73.4%). The 'infrequent' class had the lowest proportion of class members with a prescription for colchicine 11.7%, compared with the 46.3% and 50% in the 'frequent and persistent' and 'frequent then improving' classes respectively. The highest proportion of class members with a prescription for NSAIDs and a prescription for diuretics in the two years prior to baseline were seen in the 'moderately frequent then improving' (70.6%) and the 'frequent then improving' (42.9%) classes respectively.

The 'infrequent' class had the highest proportion of class members reporting taking allopurinol at each time-point as a percentage of those who returned a questionnaire (Figure 6.19). The proportion of class members taking allopurinol in the 'frequent then improving' class increased over time, reaching 8 (73%) at 36 months.

Table 6.8 Class members self-reporting allopurinol use at baseline or with prescriptions for allopurinol, colchicine, NSAIDS and diuretics in medical records in the 2 years prior to baseline

Baseline	**All in LCGA** n=1164	**'frequent and persistent'** n=95	**'gradually worsening'** n=276	**'frequent then improving'** n=14	**'moderately frequent'** n=287	**'moderately frequent then improving'** n=143	**'infrequent'** n=349
Using allopurinol*	624(53.6)	46(48.4)	129(46.7)	4(28.6)	115(40.1)	75(52.4)	255(73.0)
Allopurinol dose*							
≤100 mg	224(19.2)	25(26.3)	46(16.6)	1(7.1)	47(16.4)	29(20.3)	76(21.8)
200 mg	61(5.15)	4(4.2)	14(5.0)	0 (0.0)	10(3.5)	6(4.2)	27(7.7)
300 mg	300(25.8)	16(16.8)	52(18.8)	3(21.4)	55(19.4)	32(22.4)	142(40.7)
≥400 mg	30(2.6)	5(5.3)	14(5.1)	0 (0.0)	2(0.7)	3(2.1)	7(2.0)
Allopurinol prescription◇	635(54.6)	47(49.5)	133(48.2)	7(50.0)	118(41.1)	74(51.7)	256(73.4)
Colchicine Prescription◇	340(29.2)	44(46.3)	97(35.1)	7(50.0)	103(35.9)	48(33.6)	41(11.7)
NSAIDs Prescription◇	625(53.7)	59(62.1)	163(59.1)	9(64.3)	158(55.1)	101(70.6)	135(38.7)
Diuretics Prescription◇	280(24.1)	28(29.5)	74(26.8)	6(42.9)	67(23.3)	41(28.7)	64(18.3)

Values are n(%) unless stated otherwise

*__Self-reported allopurinol__ use and dose is taken from baseline questionnaire responses; ◇ __In medical record__ in the two years prior to baseline

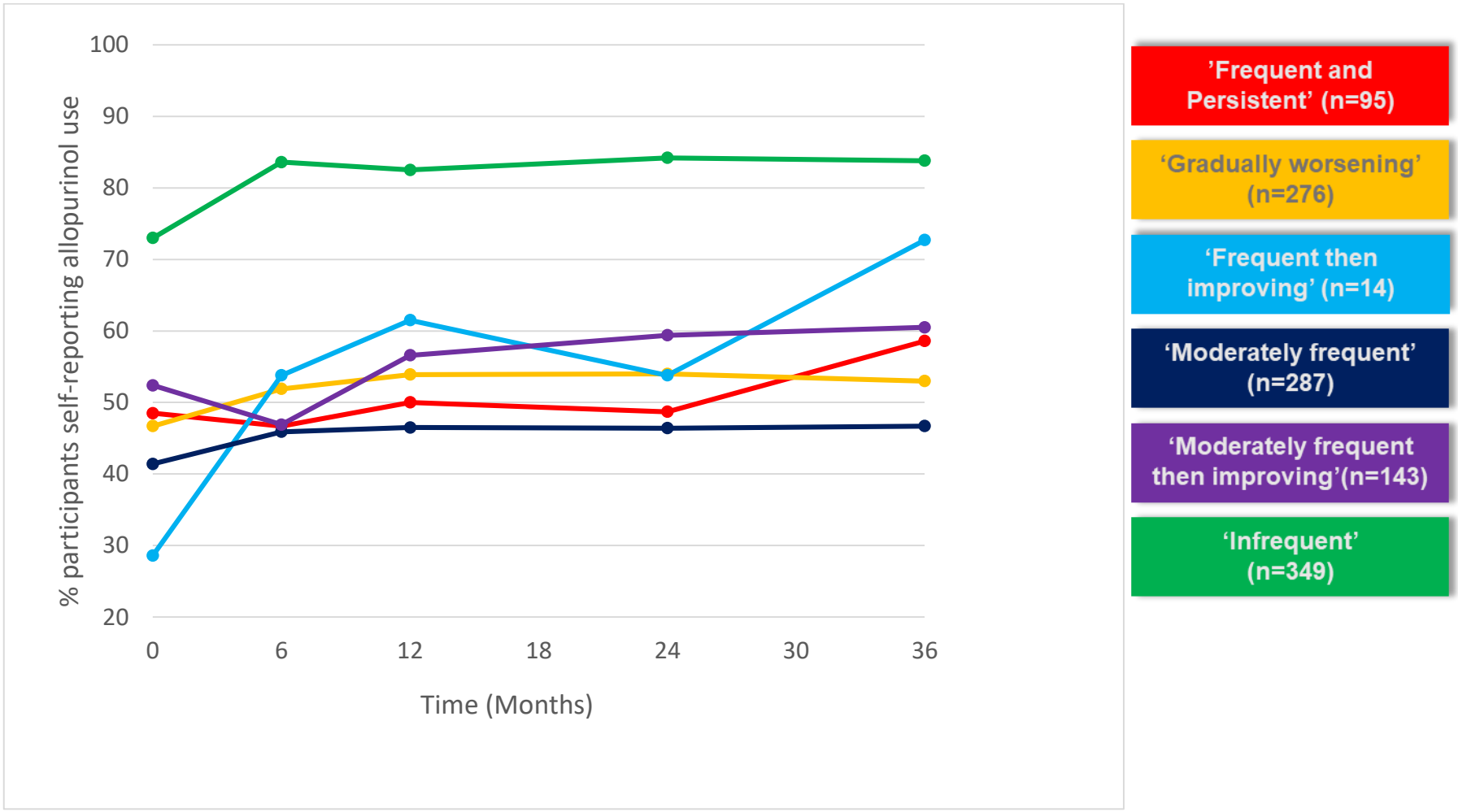


Figure 6.19 Plot of the participants per class who self-reported allopurinol use at each time-point as a percentage of those who returned a questionnaire at each time-point

Comorbidities

The most common comorbidity reported by class members in all classes was hypertension, with 61.8% of all participants reporting this comorbidity (Table 6.9). The proportion of participants with hypertension was similar across classes except for the 'frequent then improving' class who had a higher proportion of class members (78.6%) with hypertension. The 'frequent then improving', 'frequent and persistent' and 'gradually worsening' classes had higher proportions of class members with an eGFR<60 mL/min/1.73m^2 (42.9%, 30.5%, 30.1% respectively) compared with other classes. The 'frequent then improving' (14.3%), 'frequent and persistent' (9.5%) and 'gradually worsening' (8.3%) classes had higher proportions of class members with renal calculi compared with other classes. The 'frequent and persistent' and 'gradually worsening' classes had the highest proportion of class members reporting a history of myocardial infarction (MI) (11.6% and 11.6% respectively), compared with other classes. The 'frequent and persistent' class also had more class members reporting diabetes (23.2%) or a cerebrovascular accident (CVA) (6.3%) in contrast with the other five classes. The 'frequent and persistent' class had the highest mean number of self-reported comorbidities per class of 1.9, with the 'frequent then improving', 'gradually worsening', 'moderately frequent', 'moderately frequent then improving' and 'infrequent' classes having a mean number of 1.4, 1.7, 1.5, 1.5 and 1.7 respectively.

Table 6.9 Comorbidities of class members at baseline

Baseline	**All in LCGA** n=1164	**'frequent and persistent'** n=95	**'gradually worsening'** n=276	**'frequent then improving'** n=14	**'moderately frequent'** n=287	**'moderately frequent then improving'** n=143	**'infrequent'** n=349
Self-reported comorbidity							
Diabetes	202(17.4)	22(23.2)	41(14.9)	2(14.3)	48(16.7)	15(10.5)	74(21.2)
Hypertension (HT)	719(61.8)	62(65.3)	173(62.7)	11(78.6)	170(59.2)	86(60.1)	217(62.2)
Cerebrovascular Accident (CVA)	36(3.1)	6(6.3)	9(3.3)	0 (0.0)	8(2.8)	1(0.7)	12(3.4)
Transient ischaemic attack (TIA)	60(5.2)	6(6.3)	13(4.7)	0 (0.0)	12(4.2)	5(3.5)	24(6.9)
Hyperlipidaemia (HL)	499(42.9)	44(46.3)	123(44.6)	4(28.6)	116(40.4)	61(42.7)	151(43.3)
Myocardial Infarction (MI)	114(9.8)	11(11.6)	32(11.6)	1(7.1)	27(9.4)	14(9.8)	29(8.3)
Renal failure (RF)	54(4.6)	8(8.4)	19(6.9)	0 (0.0)	12(4.2)	5(3.5)	10(2.9)
Renal calculi	76(6.5)	9(9.5)	23(8.3)	2(14.3)	13(4.5)	6(4.2)	23(6.6)
Angina	142(12.2)	13(13.7)	40(14.5)	0 (0.0)	28(9.8)	19(13.3)	42(12.0)
Total comorbidities mean (SD)†	1.6(1.4)	1.9(1.6)	1.7(1.5)	1.4(0.9)	1.5(1.4)	1.5(1.3)	1.7(1.3)
eGFR <60 mL/min/1.73m^2◇	311(26.7)	29(30.5)	83(30.1)	6(42.9)	71(24.7)	34(23.8)	88(25.2)

Values are n(%) unless stated otherwise

†**Total number of comorbidities** self-reported in baseline questionnaire (diabetes, hypertension, hyperlipidaemia, myocardial infarction, angina, cerebrovascular accident, transient ischaemic attack, renal failure, renal calculi); **eGFR <60** mL/min/1.73m^2 indicative of chronic kidney disease; ◇ **In medical record** in the two years prior to baseline

Socio-demographic characteristics

The 'frequent and persistent' and 'gradually worsening' classes had the highest proportion of class members, 41.1 % and 40.6 % respectively, classified as 'most deprived' based on the IMD tertile (Table 6.10), whilst the 'frequent then improving' class had the lowest proportion (7.1%). The 'frequent and persistent' (11.6%), 'gradually worsening' (15.2%) and 'frequent then improving' (7.1%) classes had the lowest proportions of class members, who reported attending further education.

BMI and alcohol frequency

The 'frequent and persistent' class had the highest mean BMI of 30.5 kg/m^2 compared to a mean BMI ranging from 28.3 to 29.4 kg/m^2 for the other classes (Table 6.11). The 'frequent and persistent', 'gradually worsening' and 'frequent then improving' classes had the highest proportion of class members classified as obese (BMI ≥30 kg/m^2) at 43.2%, 35.1% and 35.7% respectively. The 'frequent and persistent' class had the highest proportion of class members, 17.9%, who reported never consuming alcohol and the lowest proportion of class members, 15.8%, who reported consuming alcohol daily or almost daily.

Table 6.10 Socio-demographic characteristics of class members at baseline

Baseline	**All in LCGA** n=1164	**'frequent and persistent'** n=95	**'gradually worsening'** n=276	**'frequent then improving'** n=14	**'moderately frequent'** n=287	**'moderately frequent then improving'** n=143	**'infrequent'** n=349
Male n(%)	972(83.5)	79(83.2)	226(81.9)	11(78.6)	248(86.4)	115(80.4)	293(84.0)
Age mean (SD)	65.6(12.5)	64.2(12.5)	64.5(13.1)	68.7(9.5)	65.3(12.7)	65.5(12.5)	67.0(11.7)
Age categories n(%)							
<40	25(2.2)	2(2.2)	9(3.3)	0 (0.0)	7(2.4)	3(2.1)	4(1.1)
40-49.9	114(9.8)	13(13.7)	29(10.5)	0 (0.0)	30(10.5)	17(11.9)	25(7.2)
50-59.9	207(17.8)	15(15.8)	61(22.1)	2(14.3)	55(19.2)	20(14)	54(15.5)
60-69.9	339(29.1)	32(33.7)	70(25.4)	6(42.9)	84(29.3)	45(31.5)	102(29.2)
70-79.9	333(28.6)	25(26.3)	72(26.1)	5(35.7)	72(25.1)	43(30.1)	116(33.2)
>80	146(12.5)	8(8.4)	35(12.7)	1(7.1)	39(13.6)	15(10.5)	48(13.8)
Indices of Multiple deprivation tertile n(%)							
Most deprived	360(30.9)	39(41.1)	112(40.6)	1(7.1)	82(28.6)	44(30.8)	82(23.5)
Middle Deprived	398(34.2)	29(30.5)	80(29.0)	7(50.0)	100(34.8)	51(35.7)	131(37.5)
Least deprived	406(34.9)	27(28.4)	84(30.4)	6(42.9)	105(36.6)	48(33.6)	136(39.0)
Ethnic origin n(%)							
White UK/European	1107(95.1)	93(97.9)	260(94.2)	14(100)	270(94.1)	138(96.5)	332(95.1)
Asian	16(1.4)	1(1.1)	8(2.9)	0 (0.0)	4(1.4)	1(0.7)	2(0.6)
Afro Caribbean	2(0.2)	0 (0.0)	0 (0.0)	0 (0.0)	1(.03)	0 (0.0)	1(0.3)
African	2(0.2)	0 (0.0)	1(0.4)	0 (0.0)	0 (0.0)	1(0.7)	0 (0.0)
Chinese	1(0.1)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1(0.3)
Other	7(0.6)	0 (0.0)	3(1.1)	0 (0.0)	1(0.3)	1(0.7)	2(0.6)

Table 6.10 cont. Socio-demographic characteristics of class members at baseline

Relationship status n(%)							
Married	793(68.1)	63(66.3)	192(69.6)	8(57.1)	188(65.5)	97(67.8)	245(70.2)
Widowed	112(9.6)	10(10.5)	24(8.7)	2(14.3)	33(11.5)	11(7.7)	32(9.2)
Cohabiting	73(6.3)	4(4.2)	13(4.7)	2(14.3)	24(8.4)	9(6.3)	23(6.6.)
Divorced	69(5.9)	4(4.2)	22(8.0)	1(7.1)	10(3.5)	12(8.4)	19(5.4)
Separated	22(1.9)	2(2.1)	6(2.2)		5(1.7)	3(2.1)	6(1.7)
Single	76(6.5)	11(11.6)	17(6.2)	1(7.1)	21(7.3)	9(6.3)	17(4.9)
Further education n(%)	245(21)	11(11.6)	42(15.2)	1(7.1)	78(27.2)	27(18.9)	86(24.6)

Table 6.11 BMI and alcohol frequency of class members at baseline

Baseline	**All in LCGA** n=1164	**'frequent and persistent'** n=95	**'gradually worsening'** n=276	**'frequent then improving'** n=14	**'moderately frequent'** n=287	**'moderately frequent then improving'** n=143	**'infrequent'** n=349
BMI kg/m^2 mean (SD)	29.1(5.1)	30.5(5.5)	29.4(5.0)	28.3(3.1)	28.8(4.8)	29.1(4.9)	28.8(5.3)
BMI categories n(%)							
<18.5 kg/m^2	1(0.1)	0 (0.0)	0 (0.0)	0 (0.0)	1(0.3)	0 (0.0)	0 (0.0)
18.5-24.9 kg/m^2	216(18.6)	12(12.6)	44(15.9)	1(7.1)	57(19.9)	29(20.3)	73(20.9)
25-29.9 kg/m^2	503(43.2)	36(37.9)	117(42.4)	8(57.1)	129(44.9)	60(42.0)	153(43.8)
30-34.9 kg/m^2	255(21.9)	23(24.1)	68(24.6)	5(35.7)	58(20.2)	31(21.7)	70(20.1)
35-39.9 kg/m^2	89(7.6)	12(12.6)	21(7.6)	0 (0.0)	20(7.0)	12(8.4)	23(6.6)
≥40 kg/m^2	36(3.1)	6(6.3)	8(2.9)	0 (0.0)	6(2.1)	4(2.8)	12(3.4)
Classified as obese n(%)	380(32.6)	41(43.2)	97(35.1)	5(35.7)	85(29.6)	47(32.9)	105(30.1)
Alcohol frequency n(%)							
Never	112(9.6)	17(17.9)	29(10.5)	1(7.1)	21(7.3)	11(7.7)	33(9.5)
Special occasions	154(13.2)	14(14.7)	43(15.6)	3(21.4)	31(10.8)	20(14.0)	43(12.3)
1 to 3 times per month	109(9.4)	7(7.4)	28(10.1)	1(7.1)	28(9.8)	11(7.7)	34(9.7)
Once or twice per week	249(21.4)	21(22.1)	62(22.5)	4(28.6)	61(21.3)	30(21.0)	71(20.3)
3 to 4 times per week	255(21.9)	19(20.0)	57(20.7)	2(14.3)	68(23.7)	26(18.2)	83(23.8)
Daily or almost daily	269(23.1)	15(15.8)	55(19.9)	3(21.4)	73(25.4)	44(30.8)	79(22.6)

6.4.5 Feedback from healthcare professionals and PPIE group members

The trajectory classes were presented and discussed with rheumatologists, general practitioners and other healthcare professionals who commented that the flare trajectories reflected the type of patients encountered in clinical practice (those experiencing worsening, improving and no change in symptoms) and the characteristics of class members was in keeping with the pattern of gout flares. Several healthcare professionals commented that the smallest trajectory class displayed a very interesting pattern which, as it included participants with frequent then improving flares, was of clinical relevance. Keele RUG PPIE group members were able to identify with trajectories from the different classes and PPIE group members identified classes which they felt they would be assigned to based on their own symptoms.

6.4.6 Growth mixture modelling

In addition to modelling the gout flare data using LCGA, GMM was also undertaken. The growth mixture models took over two days to run and the quadratic slope had to be fixed to zero in order for the model to run. Consequently, only the intercept parameter was permitted to vary in this model. In addition, the smaller dataset, consisting of participants who self-reported gout flare data at five time-points, needed to be used to enable the model to run. GMMs beyond two classes displayed warnings including a warning that the bootstraps for the BLRT did not converge. The lowest BIC was achieved for the two-class model; which consisted of a class with an improving trajectory and a worsening trajectory. Thus, the GMM model indicated that an improving and worsening trajectory existed within this cohort. However, the limitations described above prevented GMM being used as the primary model for selecting the optimal number of classes. The model fit indices for the GMM and the plots for the GMM can be found in appendix 13.

6.5 Discussion

6.5.1 Gout flare trajectory classes

Using LCGA, six distinct gout flare trajectory classes were identified in this primary care cohort. The selection of a six trajectories solution was based on a combination of the statistical indices and the clinical interpretation of the trajectories. This is the first-time gout flare trajectories have been described thus providing a unique insight into the occurrence of gout flares over time.

The six-class model returned both the lowest BIC in comparison to other classes and also returned a statistically significant BLRT thus indicating a better fitting model (Nylund, Asparouhov & Muthen, 2007; Peugh & Fan, 2012). The BIC and the BLRT have been identified as performing better in comparison to other indices when selecting the optimal number of latent classes (Nylund, Asparouhov & Muthen, 2007). The sensitivity analysis undertaken provided additional justification for the selection of the six-class model. In all three analyses (varying according to the number of time-points flare data were reported), the six-class model returned the lowest BIC and a statistically significant result for the BLRT. Furthermore, the trajectory plots for the six-class models in each dataset followed very similar patterns, indicating that these trajectory patterns existed in databases with varying degrees of missing gout flare data.

The six classes were allocated descriptive names to describe the classes' trajectory of gout flares; 'frequent and persistent', 'gradually worsening', 'frequent then improving', 'moderately frequent' 'moderately frequent then improving' and 'infrequent'. Thus, this six-class model displayed a range of distinct gout flare trajectories experienced by participants. In the absence of previous latent trajectory modelling of gout flares this was an exploratory analysis, however the range of gout flare trajectories identified share features with

trajectories of disease activity in spondyloarthritis (Molto et al 2017) and joint pain (Nicholls et al 2014; Verklei et al 2012) as they described classes displaying persistent, worsening, improving, moderate, and infrequent patterns. These gout flare trajectories have been presented and discussed with patients, health care professionals, and researchers, demonstrating face and content validity of the six-class model.

6.5.2 Gout flare trajectory class characteristics

The class members of the gout flare trajectory classes identified in this chapter displayed distinct gout-specific, comorbid, socio-demographic and other characteristics.

Most members in the frequent or worsening trajectory classes had experienced two or more gout flares in the 12 months prior to baseline. These classes also had higher prevalence of tophi, elevated serum urate levels and flares involving more than one joint. These findings are consistent with previous studies where tophi and having a serum urate level above 360 µmol/L were associated with a higher number of gout flares in the last year (Annemans et al 2008; Halpern et al 2009; Khanna et al 2012c; Jackson et al 2015; Rashid et al 2015; Sarawate et al 2006; Wu et al 2009).

The 'infrequent' class, where no members had experienced two or more gout flares in the previous 12 months, had more members reporting allopurinol use, the lowest mean serum urate level, and the lowest proportion of class members with oligo/polyarticular flares or tophi. The lowest proportion of participants reporting taking allopurinol at baseline was seen in the 'frequent then improving' class. This class was the only class which reported an increase in allopurinol use at subsequent time points. The lower frequency of gout flares, when the proportion of participants using ULT increases, has been described by other authors (Rees, Jenkins & Doherty, 2013). This class also had the highest proportion of class members with a prescription for diuretics in their medical records in the two years prior to baseline (42.9%)

consistent with the findings of Hunter et al (2006) that recent use of diuretics is a risk factor for gout flares (Hunter et al 2006).

The 'frequent and persistent', 'gradually worsening' and 'frequent then improving' classes had higher proportions of class members with an eGFR<60 mL/min/1.73m^2. The 'frequent and persistent' and 'gradually worsening' classes also had more participants with a self-reported history of myocardial infarction. The 'frequent and persistent' class had the highest mean BMI and the highest proportion of class members classified as obese at baseline. The 'frequent then improving' and the 'gradually worsening' classes also had a higher proportion of class members classified as obese at baseline compared to other classes. These findings are consistent with a study in the UK THIN database, which found that renal disease, ischaemic heart disease, or obesity were independent risk factors for first post-diagnosis gout flare (Rothenbacher et al 2011). The 'frequent and persistent' class also had the highest mean number of comorbidities of 1.9. An increased risk of gout flares in the presence of three or more comorbidities has been reported in other cohorts (Rashid et al 2015).

More members in the 'frequent and persistent' and 'gradually worsening' classes were classified as 'most deprived' and less had attended higher education, compared with other classes. The level of education attainment, combined with a measure of deprivation, has been advocated as a measure of socio-economic status (Grundy & Holt, 2001). Thus, these trajectory findings are consistent with previous studies which have highlighted an association between poorer health outcomes and social vulnerability & low socio-economic status (Baker, Mead & Campbell, 2002; Jaffe et al 2005). A cross-sectional analysis using baseline data from this cohort has previously shown that both deprivation and non-attendance at further education are associated with more frequent flares (Bowen-Davies et al 2018). Lower levels of educational attainment, such as no further education, have also been associated with having limited (low or marginal) health literacy (Protheroe et al 2017) and limited health

literacy has been associated with poor concordance with medication and poor overall health status (Berkman et al 2011).

6.5.3 Strengths and limitations

This LCGA was undertaken in a large cohort of 1164 participants using robust statistical methodology and optimal number of trajectory classes was determined based on the results of tests which are considered the better performing statistical tests for latent class model selection. The clinical interpretation of the trajectories, along with the characteristics of class members, provide validity to the selected model.

A range of statistical indices were considered when deciding the optimal number of latent classes, due to the absence of one commonly accepted criterion (Nylund, Asparouhov & Muthen, 2007). The result of one of the statistical indices, the LMR-LRT, returned a non-significant result for the six latent class model. However, this statistical index has been identified as being inconsistent when deciding the optimal number of classes (Jeffries, 2003; van de Schoot et al 2017) and the better performing index, the BLRT, did return a statistically significant result for the six-class model. The average posterior probability of one of the six classes in the latent class growth analysis model, the 'frequent and persistent' class, was 0.644 which is just below the level of >0.7 recommended when deciding the optimal number of latent classes (Andruff et al 2009). However, M*plus* uses observed data to supplement missing outcome responses via FIML when producing estimates (Little et al 2014) and the posterior probability result of 0.644 was returned for the class with the highest proportion of missing data and more missing outcome responses supplemented. All other classes in this six-class model returned posterior probabilities above 0.7 in the analysis in the data set of participants who responded on at least one time-point. Average posterior probabilities of >0.8 were returned for all six classes in the six-class LCGA undertaken in the two sensitivity analyses in datasets which had less or no missing data.

As this chapter reported a range of different model selection criteria, it is important to acknowledge that there is no consensus regarding which of these multiple indices should be prioritised when deciding the optimal number of latent classes in LCGA. Future research could investigate what weight should be assigned to each criterion when selecting the ideal number of classes and the application of a weighting system would help to provide uniformity in latent trajectory reporting. The current use of multiple indices without weighting can lead to challenges where indices are not in agreement and also the potential for the 'cherry picking' of indices by researchers when selecting models (van de Schoot et al 2017).

Only 14 individuals were assigned to the 'frequent then improving' class. This class size is below the level of 5%, which has been advocated by some as the ideal size of classes (Jung & Wickrama, 2008; Wickrama et al 2016). However, this class size comprised 1.2% of all participants included in LCGA and is thus above the absolute minimum class size of 1% advocated (Jung & Wickrama, 2008; Wickrama et al 2016). This class also displayed a very distinctive and clinically relevant 'improving' gout flare trajectory. Due to the exploratory nature of this analysis, such a distinct and clinically relevant trajectory could not be disregarded. However, due to the smaller class size, caution is prudent when considering the results of the subgroup analysis of the class' characteristics.

The fact that 30% of the cohort had serum urate levels above 360 μmol/L is consistent with well-documented sub optimal management of gout (Kuo et al 2015a; Roddy et al 2007b). This proportion could potentially be higher as 60% of cohort participants did not have a serum urate level within the medical record review two years prior to baseline. Suboptimal serum urate monitoring has been demonstrated in other cohorts (Wall et al 2010). The 'infrequent' class had the highest amount of missing serum urate data, which could be in part attributable to the long disease duration and the fact that serum urate may have been checked less often in these patients as their symptoms were not frequent.

6.5.4 Implications for clinical practice and further research

This identification of gout flare latent classes increases our understanding of how flares occur over time. A range of flare trajectories with distinct class characteristics were identified. This finding is of particular relevance to clinical practice as the cohort was based in primary care where the majority of patients with gout are managed.

The analysis in this chapter identified an 'infrequent' class with more class members taking allopurinol and a 'frequent then improving' class where there was an increase over time in members taking allopurinol. A reduction in flare frequency, where there is greater allopurinol use and achievement of serum urate levels, was demonstrated in a recent nurse-led intervention treat to target RCT (Doherty et al 2018).

Whilst this analysis was undertaken in an observational cohort and causal interpretations should not be surmised, the findings are compatible with existing clinical guidelines (Hui et al 2017) which advocate that urate-lowering therapy should be advised, in order to optimally manage gout symptoms, in patients who have experienced two or more gout flares in the previous 12 months.

The characteristics of members of the classes with persistently frequent or worsening gout flare trajectories may provide an opportunity to target interventions according to gout, comorbid, socio-demographic or other characteristics and subsequently improve patient care. For example, the observation that the 'frequent & persistent' and 'gradually worsening' classes all had a lower proportion of people who attended further education and a greater proportion classified as 'most deprived' could have implications for more tailored patient education and prioritisation of resources.

It is important to consider that the latent class growth analysis undertaken was intentionally exploratory in nature and the investigation of the characteristics of class members was

descriptive. As this is the first investigation of latent gout flare trajectories, it would be prudent to explore whether similar trajectories can be identified in other cohorts of people living with gout and undertake further analysis to identify potential predictors of gout flare trajectory membership. Due to the exploratory nature of this analysis a model to predict trajectory membership was not fitted.

6.6 Conclusion

In conclusion, this chapter presents the LCGA of self-reported gout flares in a prospective cohort in primary care. Six distinct gout flare trajectory classes were identified; revealing a variety of gout flare patterns over time. The classes were assigned descriptive names based on the probability of class members experiencing two or more gout flares at each time-point; 'frequent and persistent', 'gradually worsening', 'frequent then improving', 'moderately frequent', 'moderately frequent then improving' and 'infrequent'. The key characteristics of gout flare trajectories have been presented; revealing distinct characteristics specific to gout flare trajectory classes. The selection of six gout flare trajectory classes was justified by the results of statistical indices and clinical relevance. The potential clinical relevance of these gout flare trajectory classes is discussed.

In view of the range of trajectories identified, and the previously established association between gout flare frequency and health-related quality of life (HRQOL) in cross-sectional studies, the next two chapters will explore change in HRQOL over a three-year period HRQOL in people living with gout in primary care in this cohort and the factors associated with change in HRQOL.

7 Chapter Seven Change in Health-Related Quality of Life

7.1 Overview of chapter and aim

The previous chapter identified and described distinct latent classes of gout flare trajectories over a three-year period in people living with gout in primary care. Since previous studies have suggested that frequent flares are associated with poor HRQOL (see chapter one, section 1.9.4), the focus of the remainder of this thesis will be on how HRQOL changes in gout over time and which factors are associated with change in HRQOL. The aim of this chapter is to describe change in health-related quality of life (HRQOL) over a three-year period in people living with gout in primary care.

7.2 Method

7.2.1 Aim and Objectives

Aim

To describe how gout-specific and generic HRQOL in people living with gout changes over a three-year period

Objectives

1. Describe change in GIS subscales (CO, MSE, UTN, WBDA, CDA), SF-36 PF10 and HAQ-DI scores over a three-year period.
2. Describe the correlation, and distribution of the GIS subscales (CO, MSE, UTN, WBDA, CDA), SF-36 PF10 and HAQ-DI scores over a three-year period.
3. Describe the GIS subscale, SF-36 PF10 and HAQ-DI scores over three years for participants in each distinct latent gout flare trajectory class identified in chapter 6.

7.2.2 Data source

This secondary analysis used data from the study described in chapter three, including questionnaire data from baseline, 6, 12, 24 and 36-month follow-up.

7.2.3 Analysis plan

A range of descriptive analyses were used to analyse change in HRQOL scores, these have been advocated in order to investigate the wide range of potential change which can occur in longitudinal data (Leffondre et al 2004; Long, 2012; van Belle et al 2004). Longitudinal data allow both individual change and the overall mean change in an outcome to be investigated (Fayers & Machin, 2016; Van Belle et al 2004). However, as repeated measurements on the same individual are likely to be correlated (Diggle, Liang & Zeger, 1994; Long, 2012; Van Belle et al 2004) longitudinal data may not be independent (Fayers & Machin, 2016). The investigation of the extent of correlation in longitudinal data has been advocated as it can indicate the strength of the dependency in the outcome measures over time (Van Belle et al 2004; Long, 2012) and influences the way longitudinal data are statistically modelled (Fayers & Machin, 2016).

The following analysis was undertaken for the scores of the GIS subscales, SF36 PF10, and HAQ-DI, using SPSS version 24. The interpretation of GIS subscales, SF36 PF10, and HAQ-DI scores are described in chapter three, section 3.6.2 and section 3.6.3.

Measures of central tendency and spread at five time-points

A mean score (with SD) and median score (with IQR) for the GIS subscales, SF-36 PF10 and HAQ-DI were calculated for each time-point (baseline, 6, 12, 24 and 36 months).

Measures of change between time-points

The mean change (with SD) between each GIS subscale, SF-36 PF10 and HAQ-DI score at baseline and 36 months was calculated.

The standardised response mean (SRM) between baseline and 36 months was calculated by dividing the result of the mean change, by the standard deviation of change scores (Fayers & Machin, 2016; Morris, 2000; Morris & DeShon, 2002; Wolff Smith & Beretvas, 2009). This method has been recommended as a measurement of effect size in repeated measures data (Fayers & Machin, 2016; Morris, 2000; Morris & DeShon, 2002; Wolff Smith & Beretvas, 2009).

The effect size between baseline and 36 months was also calculated for the HRQOL measures by dividing the result of the mean change, by the standard deviation of baseline scores (Fayers & Machin, 2016; Middel & Van Sonderen, 2002; Sim & Wright, 2000).

Mean change (SD), SRM and effect size were also calculated between baseline and 12 months to allow comparison with existing studies.

Graphing longitudinal data

Graphing longitudinal data has been advocated to provide information about mean change but also to indicate the degree of variability and change at an individual level (Locascio & Atri, 2011; Long, 2012). The median GIS subscale, SF-36 PF10 and HAQ-DI with IQR were plotted at each time point as box plots to visualise change in each HRQOL score over time. Spaghetti plots were created showing plots of the pattern of change in the HRQOL outcome measure for each individual and the individual variability within the cohort (Locascio & Atri, 2011; Long, 2012; van Belle et al 2004), thus demonstrating within-subject change and between-subject differences in change (Locascio & Atri, 2011). An interpolation line was fitted to the spaghetti plots to indicate the overall pattern of change.

Correlation and distribution of outcome measures

To investigate the correlation between outcome measures at different time-points, and hence the level of dependency, scatterplots were created plotting the HRQOL outcome at one time-point with another time-point, thus creating a matrix of scatter plots with related correlation

coefficients. Histograms and quantile-quantile plots (Q-Q plots) of the GIS subscales, PF10 and HAQ-DI were undertaken in order to determine whether the scores were normally distributed (Bland, 2015; Ghasemi & Zahediasl, 2012) with the Q-Q plots used to compare the quantiles of the observed frequency distribution against the corresponding quantiles of the theoretical normal distribution..

The investigation of the linear form, correlation, and distribution of these HRQOL scores was undertaken to inform how they should be statistically modelled later in the thesis.

HRQOL over three years in each latent gout flare trajectory class

Mean (SD) GIS subscale, SF-36 PF10 and HAQ-DI scores were calculated at each time-point (baseline, 6, 12, 24 and 36 months) for participants in each distinct latent gout flare trajectory class identified in the previous chapter.

7.3 Results

7.3.1 Measures of central tendency and spread of HRQOL scores over five time-points

The mean (SD) GIS subscale scores at baseline ranged from 33.46 (20.57) for the GIS UTN to 48.65 (28.33) for the GIS CO (Table 7.1). The mean (SD) GIS subscale scores at 36 months ranged from 28.74 (19.19) for the GIS UTN to 42.57 (27.85) for the GIS CO (Table 7.1). The median (IQR) GIS subscale scores at baseline ranged from 33.33 (16.67, 41.67) for the GIS UTN to 50 (25.00, 67.75) for the GIS CO (Figure 7.1). The median (IQR) GIS subscale scores at 36 months ranged from 25 (16.67, 41.67) for the GIS UTN to 37.50 (25.00, 62.50) for the GIS MSE (Figure 7.1). The median (IQR) GIS MSE and CDA subscale scores did not change over the three years (Figure 7.1). Lower mean and median GIS scores at 36 months for the GIS CO, UTN and WBDA indicate better HRQOL at a group level.

Mean (SD) SF-36 PF10 score was 75.91 (26.12) at baseline and 80.63 (24.49) at 36 months, median (IQR) scores were 90 (53.33, 96.67) at baseline and 90 (66.67, 100) at 36 months (Table

7.1 and Figure 7.2 respectively). Mean (SD) HAQ-DI score was 0.51 (0.71) at baseline and 0.47 (0.67) at 36 months, median (IQR) scores were 0.13 (0.00, 0.88) at baseline and 0.13 (0.00, 0.75) at 36 months (Table 7.1 and Figure 7.2 respectively). The median (IQR) scores for the SF-36 PF10 and HAQ-DI did not change over 36 months (Figure 7.2).

Table 7.1 Mean (SD) GIS subscales, SF-36 PF10 and HAQ-DI score at each time-point (baseline, 6, 12, 24, 36 months)

Time-point (months)	**GIS CO**		**GIS MSE**		**GIS UTN**		**GIS WBDA**		**GIS CDA**		**SF-36 PF10**		**HAQ-DI**	
	Sample	**Mean (SD)**	Sample	**Mean (SD)**	Sample	**Mean (SD)**	Sample	**Mean (SD)**	Sample	**Mean (SD)**	Sample	**Mean (SD)**	Sample	**Mean (SD)**
0	1146	48.65 (28.33)	1130	40.45 (26.33)	1108	33.46 (20.57)	1150	45.19 (26.41)	1139	40.13 (24.35)	960	75.91 (26.12)	1144	0.51 (0.71)
6	801	48.68 (27.80)	789	40.91 (26.35)	776	31.77 (20.16)	792	43.20 (25.31)	779	40.26 (23.82)	666	78.28 (25.56)	807	0.49 (0.70)
12	703	46.37 (27.27)	695	39.10 (25.57)	685	30.72 (20.13)	686	40.47 (25.88)	699	39.72 (23.91)	579	77.36 (27.29)	699	0.46 (0.68)
24	677	44.77 (27.52)	673	38.86 (27.32)	656	30.56 (19.49)	667	41.19 (26.65)	677	39.17 (24.07)	550	80.25 (24.77)	688	0.45 (0.68)
36	588	42.57 (27.85)	578	38.02 (26.20)	573	28.74 (19.19)	586	39.64 (25.49)	588	38.94 (24.03)	482	80.63 (24.49)	592	0.47 (0.67)

GIS subscales scored from 0 to 100; higher scores on each scale indicating a greater impact of gout on HRQOL/worse HRQOL.
SF-36 PF10 scored from 0 to 100; higher score indicating performs all types of physical activities including the most vigorous without limitations due to health.
HAQ-DI scored from 0 to 3; higher score indicating greater activity limitation.
Baseline HRQOL scores for cohort has been published previously (Chandratre et al 2018).

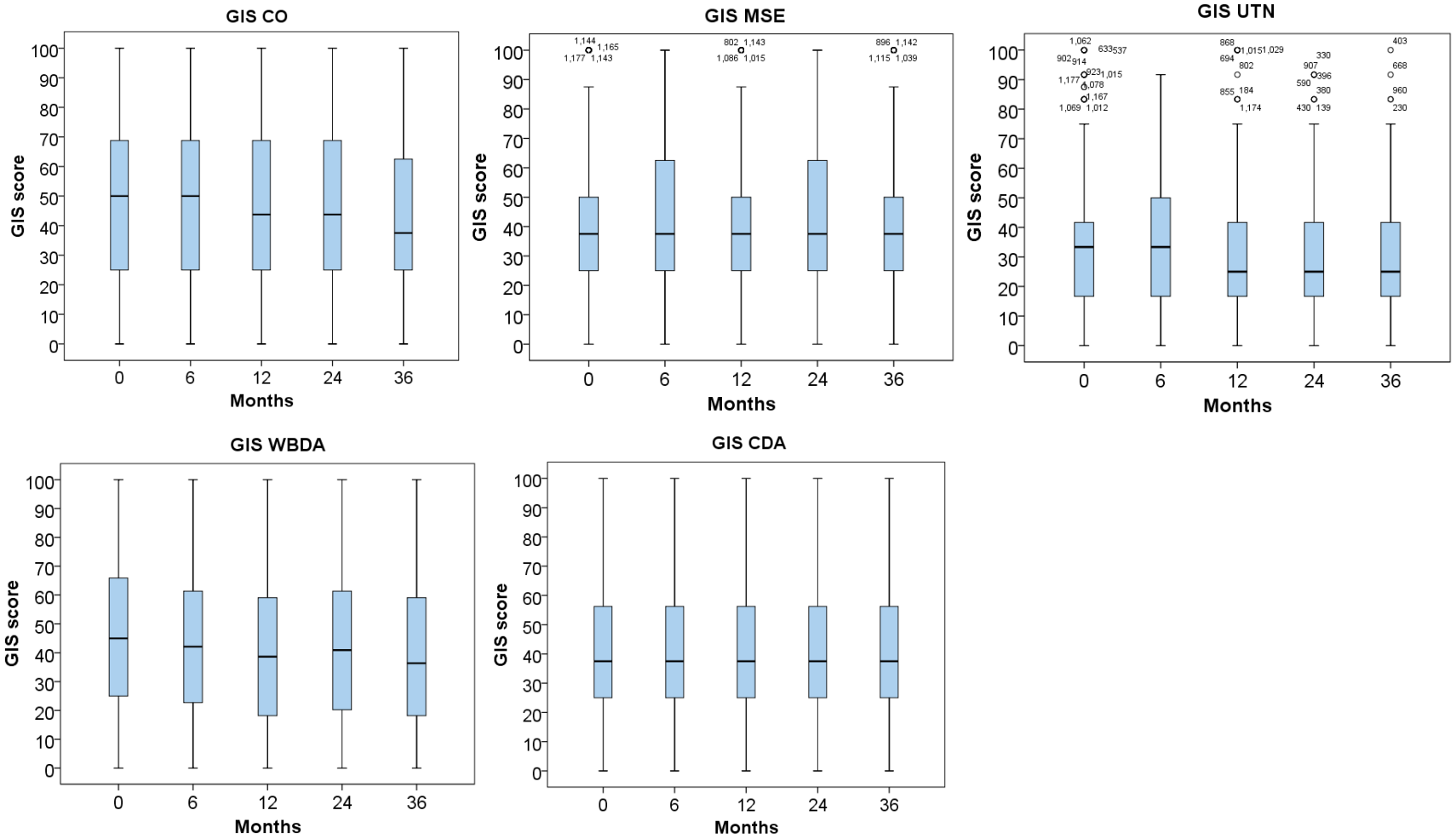


Figure 7.1 Box and whisker plot displaying the median, IQR, and outliers for the GIS subscale scores at each time-point (baseline, 6, 12, 24, 36 months)

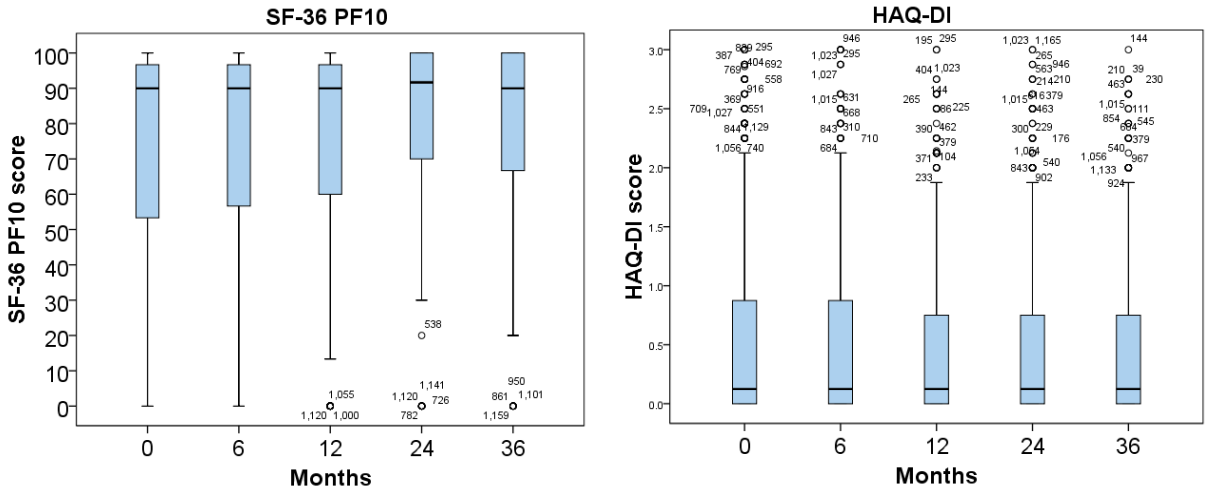


Figure 7.2 Box and whisker plot displaying the median, IQR, and outliers for the SF-36 PF10 and HAQ-DI scores at each time-point (baseline, 6, 12, 24, 36 months)

7.3.2 Mean change, standardised response mean and effect size of HRQOL scores

Baseline to 36 months

The mean change (SD) in scores from baseline to 36 months for the GIS subscales ranged from -0.46 (20.03) for the GIS CDA to -6.36 (20.31) for the GIS WBDA (Table 7.2). The SRM between baseline and 36 months ranged from -0.02 for the GIS CDA to -0.31 for the GIS WBDA and the effect sizes ranged from -0.02 for the GIS CDA to -0.24 for the GIS WBDA (Table 7.2). The SRM for the GIS WBDA and GIS CO were consistent with small effect sizes (-0.31, -0.21 respectively) (Cohen, 1992). The change in GIS WBDA score was consistent with a small effect size (-0.24) (Cohen, 1992).

The mean change (SD) between scores at baseline and 36 months for the SF-36 PF10 was -0.34 (18.87) (Table 7.2). The change in SF-36 PF10 score was consistent with a negligible effect size (0.02) (Cohen, 1992).

The mean change (SD) between scores at baseline and 36 months for the HAQ-DI was 0.08 (0.39) (Table 7.2). The SRM for HAQ-DI is consistent with a small effect size (0.21) (Cohen, 1992). The effect size of the HAQ-DI score was consistent with a negligible effect size (0.13) (Cohen, 1992).

Table 7.2 The mean scores (SD) at baseline and 36 months, mean change (SD) from baseline to 36 months, standardised response mean (SRM) and effect size (ES) for the GIS subscales, PF10 and HAQ-DI.

HRQOL measure	**Change in scores Baseline to 36 Months**					
	Sample	Mean at baseline (SD)	Mean at 36 months (SD)	Mean change [CI] (SD)	SRM◇	Effect size◇◇
GIS CO	576	47.77 (28.21)	42.64 (27.92)	**-5.13 [-7.18,-3.08]** (25.06)	-0.21	-0.18
GIS MSE	567	40.65 (26.25)	38.12 (26.31)	**-2.54 [-4.63,-0.05]** (25.19)	-0.10	-0.10
GIS UTN	556	30.48 (19.25)	28.68 (18.95)	**-1.80 [-3.57,-0.03]** (21.27)	-0.09	-0.09
GIS WBDA	579	45.97 (26.22)	39.61 (25.41)	**-6.36 [-8.1,-4.70]** (20.31)	-0.31	-0.24
GIS CDA	575	39.59 (23.83)	39.13 (23.80)	-0.46 [-2.10,1.18] (20.03)	-0.02	-0.02
SF-36 PF10	420	82.68 (23.05)	82.34 (23.54)	-0.34 [-2.15,1.47] (18.87)	-0.02	-0.02
HAQ-DI	577	0.38 (0.62)	0.46 (0.67)	**0.08 [0.05,0.11]** (0.39)	0.21	0.13

◇**Standardised response mean** (SRM) calculated as mean change/SD of change scores ◇◇**Effect size** calculated as mean change/SD of scores at baseline.

GIS subscales scored from 0 to 100; higher scores on each scale indicating a greater impact of gout on HRQOL/worse HRQOL.
SF-36 PF10 scored from 0 to 100; higher score indicating performs all types of physical activities including the most vigorous without limitations due to health.
HAQ-DI scored from 0 to 3; higher score indicating greater activity limitation.

Baseline to 12 months

Change between baseline and 12 months was analysed to allow comparison with previous published studies.

The mean change (SD) in scores from baseline to 12 months for the GIS subscales ranged from 0.20 (19.21) for the GIS CDA to -5.45 (19.11) for the GIS WBDA (Table 7.3). The SRM for the change in GIS subscales between baseline and 12 months ranged from 0.01 for the GIS CDA to -0.29 for the GIS WBDA and the effect size ranged from 0.01 for the GIS CDA to -0.21 for the GIS WBDA (Table 7.3). The change in GIS WBDA was consistent with a small effect size (0.21) (Cohen, 1992).

The mean change (SD) between SF-36 PF10 scores at baseline and 12 months was -1.46 (22.30) (Table 7.3). The change in SF-36 PF10 was consistent with a negligible effect size (0.06) (Cohen, 1992).

The mean change (SD) between scores at baseline and 12 months for the HAQ-DI was 0.03 (0.37) (see Table 7.3). The change in HAQ-DI was consistent with a negligible effect size (0.05) (Cohen, 1992).

Table 7.3 The mean scores (SD) at baseline and 12 months, mean change (SD) from baseline to 12 months, standardised response mean (SRM) and effect size (ES) for the GIS subscales, SF-36 PF10 and HAQ-DI

HRQOL measure	**Change in scores Baseline to 12 Months**					
	Sample	Mean at baseline (SD)	Mean at 12 months (SD)	Mean change [CI] (SD)	SRM◇	Effect size◇◇
GIS CO	687	47.71 (27.99)	46.37 (27.26)	-1.34 [-3.01,0.33] (22.26)	-0.06	-0.05
GIS MSE	678	40.17 (25.75)	39.14 (25.52)	-1.03 [-2.65,0.58] (21.42)	-0.05	-0.04
GIS UTN	663	30.90 (19.92)	30.50 (19.90)	-0.40 [-2.01,1.2] (21.02)	-0.02	-0.02
GIS WBDA	674	45.45 (25.68)	40.00 (25.68)	**-5.45 [-6.89,-4.00]** (19.11)	-0.29	-0.21
GIS CDA	683	39.48 (23.84)	39.68 (23.85)	0.20 [-1.23,1.65] (19.21)	0.01	0.01
SF-36 PF10	496	80.92 (24.0)	79.46 (26.57)	-1.46 [-3.43,0.51] (22.30)	-0.06	-0.06
HAQ-DI	681	0.43 (0.64)	0.46 (0.67)	0.03 [-3.43,0.51] (0.37)	0.08	0.05

◇**Standardised response mean** (SRM) calculated as mean change/SD of change scores ◇◇**Effect size** calculated as mean change/SD of scores at baseline.

GIS subscales scored from 0 to 100; higher scores on each scale indicating a greater impact of gout on HRQOL/worse HRQOL.
SF-36 PF10 scored from 0 to 100; higher score indicating performs all types of physical activities including the most vigorous without limitations due to health.
HAQ-DI scored from 0 to 3; higher score indicating greater activity limitation.

7.3.3 Spaghetti plots of individual change in HRQOL scores over five time-points

The spaghetti plots of GIS subscales show the diverse range of different GIS scores at baseline and also patterns of change in GIS scores over 36 months (Figure 7.3).

The spaghetti plot for the GIS WBDA subscale scores (Figure 7.3) shows a greater variation in individual GIS scores at baseline and individual patterns of GIS scores over time, compared to other GIS subscales. Although all subscale scores show a wide variation in individual scores at baseline and over time.

The spaghetti plot of SF-36 PF10 and HAQ-DI scores over 36 months (Figure 7.4) show the wide range of SF-36 PF10 and HAQ-DI scores at baseline and the patterns of longitudinal change in the scores. The plots display the greater proportion of higher SF-36 PF10 scores and lower HAQ-DI scores over time.

The interpolation lines fitted to all the HRQOL scores in Figure 7.3 and Figure 7.4 indicate a linear pattern to the scores over the 36-month period.

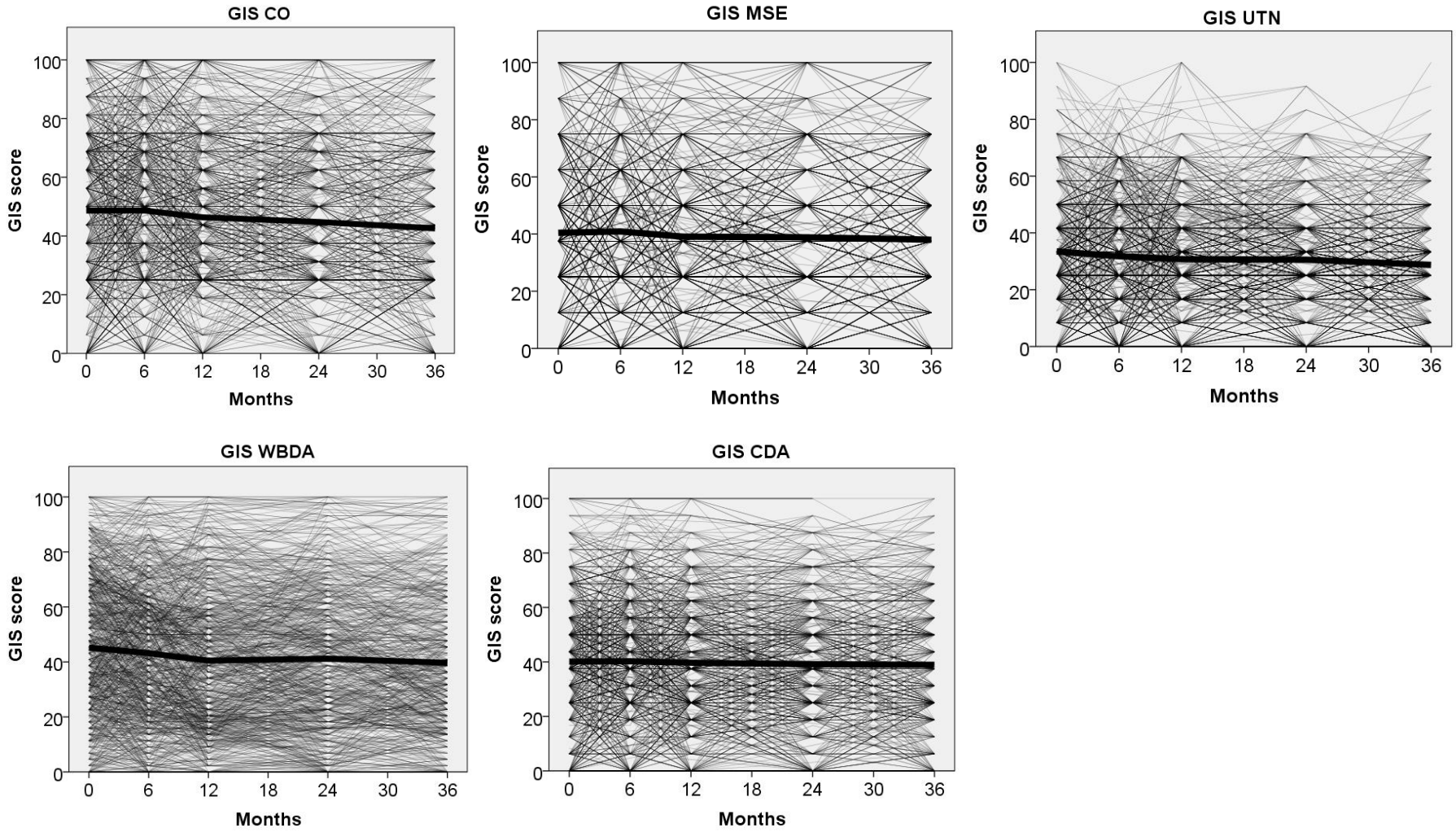


Figure 7.3 Spaghetti plots of individual participants' GIS scores over 36 months with interpolation line

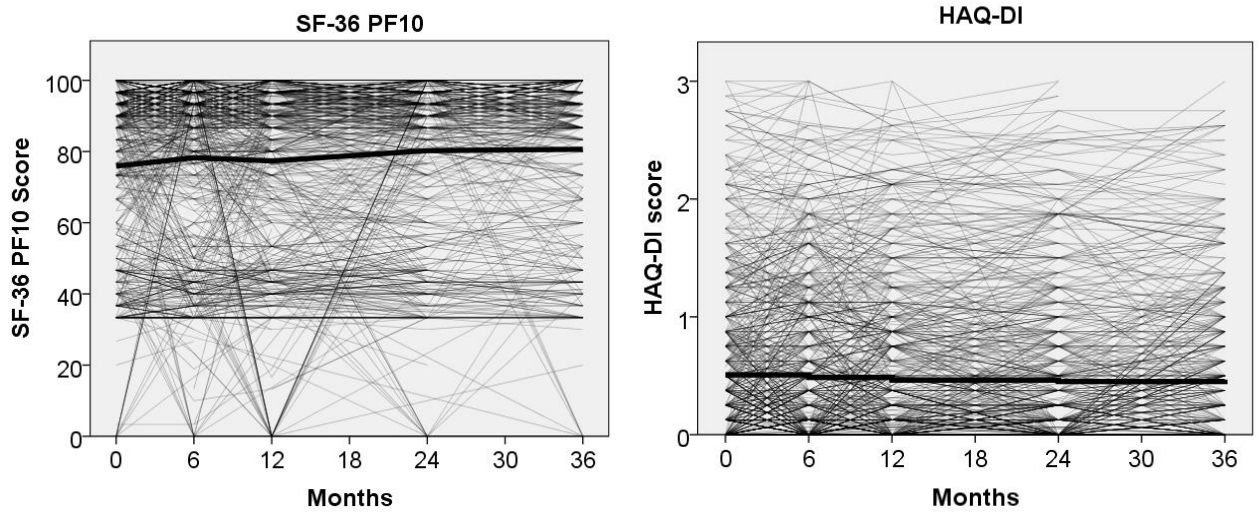


Figure 7.4 Spaghetti plot of individual participants' SF-36 PF10 and HAQ-DI scores over 36 months with interpolation line

7.3.4 Correlation and distribution of outcome measures

All the GIS score scatterplots (Figure 7.5) display a positive linear relationship between GIS scores at different time-points. The correlation coefficients, r, ranged from 0.38 (between GIS UTN scores at baseline and 36 months) displaying moderate correlation, to 0.80 (between GIS WBDA at 24 and 36 months) displaying very high correlation (Figure 7.5). The histograms of GIS subscales (Figure 7.5) display some symmetry suggestive of a normal distribution, however the number of lower scores were higher than would be expected for a perfect normal distribution. The Q-Q plots in appendix 14 confirm that the distribution of GIS scales is close to a normal distribution.

All scatterplots of SF-36 PF10 and HAQ-DI scores (Figure 7.6) display a positive correlation between scores at different time-points. The correlation coefficients for SF-36 PF10 scores ranged from 0.62 (between baseline and 12 months) and 0.77 (between 24 and 36 months) displaying a high correlation (Figure 7.6). The SF-36 PF10 scatter plots and histograms showed two sub-groups (Figure 7.6). The histograms also show a left skew; with a greater frequency of higher SF-36 PF10 scores. The histograms also show a second much smaller peak in lower scores. The correlation coefficients for HAQ-DI scores ranged from 0.82 (between baseline and 24 months, baseline and 36 months) and 0.86 (between 6 and 12 months, 12 and 24 months), indicating a very high correlation between HAQ-DI at different time-points over all (Figure 7.6). The histograms of HAQ-DI at each time-point display a right skew; with a greater frequency of lower scores. The Q-Q plots for SF-36 PF10 scores and HAQ-DI scores at each time-point in appendix 14 confirm the left and right skew in these outcomes respectively.

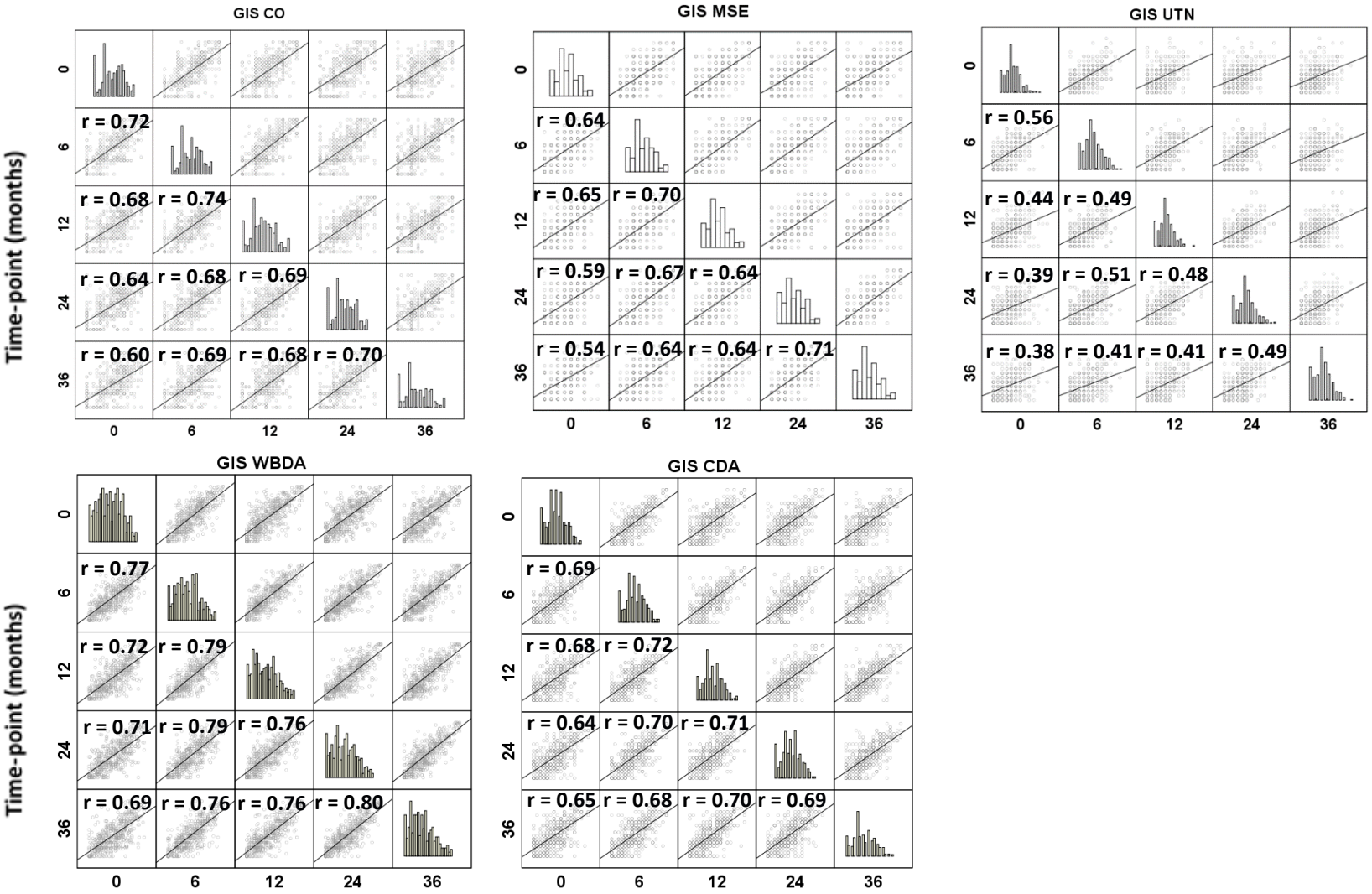


Figure 7.5 Matrices of GIS score scatterplots at different time-points with the line of best fit and correlation coefficient on each scatter plot, with histograms of GIS scores at each time-point displayed on the diagonal of the matrix

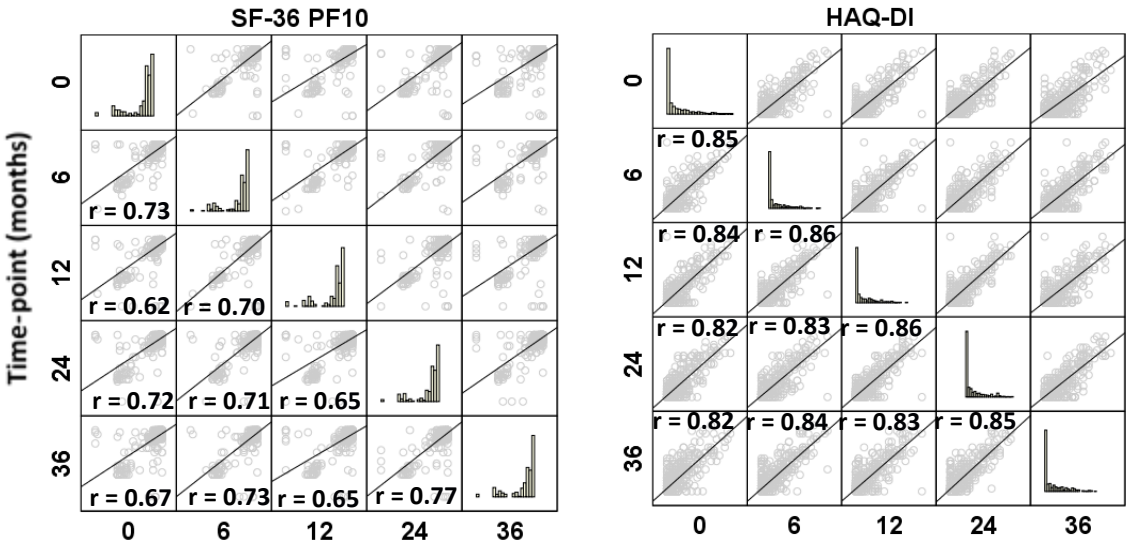


Figure 7.6 Matrices of SF-36 PF10 and HAQ-DI score scatterplots at different time-points with the line of best fit and correlation coefficient on each scatter plot, with histograms of SF-36 PF10 and HAQ-DI scores at each time-point displayed on the diagonal of each matrix

7.3.5 HRQOL over three years in each latent gout flares trajectory class

Table 7.4 to 7.10 display the mean (SD) GIS subscale, SF-36 PF10 and HAQ-DI scores at each questionnaire time-point per gout flare trajectory class.

At baseline the 'infrequent' class had the lowest score on each of the five GIS subscales, indicating better disease-specific HRQOL, compared with other classes. At each time-point the 'frequent and persistent' and 'gradually worsening' classes had higher mean scores on each of the GIS subscale scores, indicating worse disease-specific HRQOL, in comparison with other classes (Table 7.4 to Table 7.8). The greatest decrease in all GIS subscale scores from baseline to 36 months was observed in the 'frequent then improving' and 'moderately frequent then improving' classes.

At each time-point the 'frequent and persistent' and 'gradually worsening' classes had lower mean scores on the SF-36 PF10 and higher mean scores on the HAQ-DI, indicating worse generic HRQOL and greater activity limitation respectively, in comparison with other classes (Table 7.9 to Table 7.10). In the 'frequent and persistent' class the mean HAQ-DI score was over 1.0, indicating moderate functional disability, at all time-points except 24 months. The mean HAQ-DI score in the 'frequent and persistent' class at 36 months (1.22) was higher than at any previous time-point.

Table 7.4 Mean (SD) GIS CO scores per gout flare trajectory class

	GIS CO Baseline	**GIS CO** 6 months	**GIS CO** 12 months	**GIS CO** 24 months	**GIS CO** 36 months
Frequent & persistent	78.46 (20.67)	77.67 (21.54)	75.15 (22.02)	75.17 (21.42)	68.32 (30.16)
Gradually worsening	61.15 (22.29)	64.41 (23.38)	62.20 (23.17)	61.64 (24.03)	59.39 (24.72)
Frequent then improving	61.61 (22.45)	60.10 (21.88)	56.73 (27.77)	39.42 (23.16)	45.00 (28.38)
Moderately frequent	47.15 (25.98)	49.77 (24.16)	47.15 (23.18)	44.98 (25.20)	44.33 (25.10)
Moderately frequent then improving	46.33 (26.18)	41.02 (24.64)	35.98 (24.51)	36.44 (25.46	29.61 (24.07)
Infrequent	32.19 (25.88)	33.97 (26.02)	32.43 (24.99)	30.88 (23.10)	30.52 (23.99)
All in LCGA	48.55 (28.31)	48.68 (27.80)	46.37 (27.27)	44.77 (27.52)	42.57 (27.85)

Table 7.5 Mean (SD) GIS MSE scores per gout flare trajectory class

	GIS MSE Baseline	**GIS MSE** 6 months	**GIS MSE** 12 months	**GIS MSE** 24 months	**GIS MSE** 36 months
Frequent & persistent	58.10 (25.50)	60.37 (24.22)	60.94 (26.88)	62.16 (25.60)	56.48 (25.33)
Gradually worsening	49.33 (25.06)	50.62 (24.60)	49.25 (21.66)	50.51 (26.08)	49.22 (25.13)
Frequent then improving	45.54 28.0	50.96 (23.64)	45.19 (30.15)	37.50 (21.97)	40.00 (23.42)
Moderately frequent	36.37 (25.01)	39.73 (24.17)	37.76 (23.24)	36.62 (24.31)	35.78 (22.72)
Moderately frequent then improving	37.41 (23.90)	35.64 (24.61)	31.37 (23.71)	29.17 (25.76)	29.38 (26.39)
Infrequent	32.60 (25.39)	33.64 (26.97)	31.94 (26.49)	32.60 (26.70)	32.92 (26.32)
All in LCGA	40.22 (26.34)	40.91 (26.35)	39.10 (25.57)	38.86 (27.32)	38.02 (26.20)

Table 7.6 Mean (SD) GIS UTN scores per gout flare trajectory class

	GIS UTN Baseline	**GIS UTN** 6 months	**GIS UTN** 12 months	**GIS UTN** 24 months	**GIS UTN** 36 months
Frequent & persistent	48.63 (22.01)	47.96 (21.21)	51.25 (26.80)	44.52 (20.80)	45.06 (20.58)
Gradually worsening	38.15 (19.34)	39.97 (17.72)	37.33 (18.14)	39.31 (18.61)	38.64 (18.0)
Frequent then improving	50.00 (26.55)	39.74 (28.09)	36.54 (20.84)	22.92 (12.37)	25.00 (13.94)
Moderately frequent	34.75 19.27	33.41 (18.26)	33.00 (18.27)	30.68 (18.37)	28.70 (17.91)
Moderately frequent then improving	33.64 (18.68)	30.42 (19.70)	26.66 (18.09)	25.05 (18.37)	23.54 (19.31)
Infrequent	23.91 (18.23)	21.60 (17.73)	21.09 (17.0)	24.13 (17.90)	21.65 (16.63)
All in LCGA	33.46 (20.59)	31.77 (20.16)	30.72 (20.13)	30.56 (19.49)	28.74 (19.19)

Table 7.7 Mean (SD) GIS WDA scores per gout flare trajectory class

	GIS WBDA Baseline	**GIS WBDA** 6 months	**GIS WBDA** 12 months	**GIS WBDA** 24 months	**GIS WBDA** 36 months
Frequent & persistent	58.50 (25.37)	58.32 (24.64)	60.25 (26.8)	55.09 (29.41)	55.40 (26.78)
Gradually worsening	48.29 (23.57)	46.39 (22.88)	44.39 (25.31)	46.64 (26.05)	44.47 (24.04)
Frequent then improving	41.33 (21.86)	42.90 (16.56)	40.28 (24.50)	33.59 (23.23)	28.77 (14.75)
Moderately frequent	42.93 (27.02)	38.86 (23.38)	34.51 (22.77)	35.50 (24.92)	34.67 (24.06)
Moderately frequent then improving	45.01 (26.72)	42.68 (27.34)	38.64 (26.48)	39.90 (26.91)	36.22 (26.08)
Infrequent	40.89 (27.13)	42.02 (26.76)	40.41 (26.64)	40.95 (26.72)	40.3 (26.27)
All in LCGA	45.08 (26.45)	43.20 (25.31)	40.47 (25.88)	41.19 (26.65)	39.64 (25.49)

Table 7.8 Mean (SD) GIS CDA scores per gout flare trajectory class

	GIS CDA Baseline	**GIS CDA** 6 months	**GIS CDA** 12 months	**GIS CDA** 24 months	**GIS CDA** 36 months
Frequent & persistent	55.60 (28.42)	58.50 (26.27)	57.44 (31.61)	56.08 (31.57)	51.94 (29.51)
Gradually worsening	46.59 (21.73)	50.01 (22.65)	49.59 (22.24)	49.06 (24.39)	47.35 (22.87)
Frequent then improving	46.88 (18.79)	45.67 (19.83)	47.12 (27.08)	38.94 (21.82)	42.50 (12.77)
Moderately frequent	37.5 (23.15)	37.56 (20.19)	37.90 (20.13)	35.85 (20.69)	36.95 (22.33)
Moderately frequent then improving	40.18 (24.85)	38.28 (24.09)	31.98 (23.67)	35.78 (23.19)	31.33 (24.17)
Infrequent	32.20 (22.67)	32.79 (23.11)	33.99 (22.47)	33.51 (22.45)	35.76 (23.57)
All in LCGA	39.94 (24.32)	40.26 (23.82)	39.72 (23.91)	39.17 (24.07)	38.94 (24.03)

Table 7.9 Mean (SD) SF-36 PF10 scores per gout flare trajectory class

	SF-36 PF10 Baseline	**SF-36 PF10** 6 months	**SF-36 PF10** 12 months	**SF-36 PF10** 24 months	**SF-36 PF10** 36 months
Frequent & persistent	57.69 (25.27)	56.09 (25.24)	49.52 (23.67)	61.23 (24.56)	59.85 (25.31)
Gradually worsening	68.95 (27.42)	73.20 (25.32)	71.43 (26.61)	72.08 (25.45)	76.32 (24.12)
Frequent then improving	78.25 (23.78)	82.25 (18.94)	83.03 (18.59)	74.14 (32.85)	75.67 (34.75)
Moderately frequent	81.20 (23.65)	81.90 (23.77)	81.42 (24.42)	84.28 (21.69)	82.86 (23.58)
Moderately frequent then improving	79.88 (23.00)	80.23 (27.19)	79.18 (28.53)	79.36 (26.00)	85.38 (20.51)
Infrequent	80.31 (25.07)	81.56 (24.15)	81.38 (27.30)	85.69 (23.05)	82.70 (24.77)
All in LCGA	76.14 (25.98)	78.28 (25.56)	77.36 (27.29)	80.25 (24.77)	80.63 (24.49)

Table 7.10 Mean (SD) HAQ-DI scores per gout flare trajectory class

	HAQ-DI Baseline	**HAQ-DI** 6 months	**HAQ-DI** 12 months	**HAQ-DI** 24 months	**HAQ-DI** 36 months
Frequent & persistent	1.01 (0.89)	1.09 (0.81)	1.10 (0.77)	0.90 (0.85)	1.22 (0.95)
Gradually worsening	0.60 (0.73)	0.64 (0.76)	0.61 (0.70)	0.66 (0.76)	0.59 (0.68)
Frequent then improving	0.56 (0.86)	0.45 (0.65)	0.51 (0.74)	0.51 (0.76)	0.54 (0.65)
Moderately frequent	0.44 (0 .66)	0.42 (0.66)	0.41 (0.65)	0.35 (0.60)	0.38 (0.61)
Moderately frequent then improving	0.42 (0.65)	0.44 (0.66)	0.41 (0.67)	0.45 (0.69)	0.41 (0.60)
Infrequent	0.39 (0.64)	0.33 (0.60)	0.29 (0.55)	0.31 (0.56)	0.35 (0.60)
All in LCGA	0.50 (0.71)	0.49 (0.70)	0.46 (0.68)	0.45 (0.68)	0.47 (0.67)

7.4 Discussion

This chapter describes longitudinal change in gout-specific and generic HRQOL over three years in people living with gout. The variance, correlation, distribution and functional form of HRQOL scores over three years was described. The baseline GIS subscales, SF-36 PF10 and HAQ-DI scores in this cohort have been published previously (Chandratre et al 2018), thus the focus in this chapter was longitudinal change.

7.4.1 HRQOL group level change and comparison with previous studies

The measures of central tendency revealed a reduction in the GIS CO, UTN and WBDA subscale scores over the three-year period, thus indicating improving HRQOL at a group level. However, the GIS MSE and CDA scores changed negligibly over the three-year period.

The mean change (SD) in all GIS subscales between baseline to 36 months displayed better HRQOL at 36 months, however there was negligible change in the GIS CDA. Only the SRM and effect size for the WBDA and the SRM for the GIS CO was above 0.20, consistent with a small effect size. This is the first study to collect GIS subscales scores over a three-year period, hence comparison of the change in GIS subscales observed over three years with other published results was not possible. Consequently, change in HRQOL between baseline and 12 months was also analysed to allow comparison to existing literature. In comparison to mean change observed between baseline to 36 months, less change in the GIS subscales was detected between baseline and 12 months. Of all of the GIS subscales only the GIS WBDA had an SRM and effect size over 0.20 over 12 months, consistent with a small effect.

Change in GIS subscales have been reported from baseline to one year in an US longitudinal cohort by Wallace et al (2016). The change in GIS subscales scores, SRMs and effect sizes from this study also displayed a reduction in GIS subscale scores over time, indicating an improvement in HRQOL. However, smaller mean changes, SRMs and effect sizes were

observed for the GIS subscales in this cohort compared with this previous prospective study. There were also differences in the subscales displaying the most and least change in this cohort compared to previous studies. For example, in contrast with this cohort the GIS CO had the greatest mean change and the GIS MSE had the smallest mean change in the study by Wallace et al (2016). Improvement in GIS CO, UTN, WBDA and CDA scores were observed from week four to week eight in an US medication RCT by Khanna et al (2011b) in groups of participants who were deemed to have 'minimally improved' gout based on anchor questions. Mean change and effect sizes reported by Khanna et al (2011b) were larger than the effect sizes (range) observed in this cohort. Change in GIS CO and UTN scores over two years in participants living with gout was reported by Doherty et al (2018) in a treat to target RCT. Improvements in GIS CO and UTN scores were reported over the two years for both the 'nurse led' and 'usual care' groups, with a greater improvement in HRQOL observed in the nurse led group.

The measures of central tendency revealed little change in SF-36 PF10 scores overtime and the SRM and effect sizes for the change in SF-36 PF10 scores from baseline to 36 months and baseline to 12 months were all below 0.20, thus indicating a negligible change (Cohen, 1992). The measures of central tendency also displayed negligible change for the HAQ-DI over time. However, the mean change between baseline and 36 months displayed a very small increase in HAQ-DI scores indicating worse activity limitation HRQOL, and the SRM for the change in HAQ-DI scores from baseline to 36 months was 0.21, thus indicating a small change. This is the first study in gout to collect SF-36 PF10 and HAQ-DI scores over a follow-up period of longer than two years. In observational studies, effect sizes for the SF-36 PF10 were reported over 12 months in people with gout in a Spanish single centre cohort (Khanna et al 2011a) and for the HAQ-DI over 6 months in people with gout in a Mexican multicentre cohort study (Alvarez-Hernandez et al 2008). The effect sizes reported by these studies were larger than the effect sizes (range) observed in this cohort.

The differences between the magnitude of change in GIS subscales, SF-36 PF10, and HAQ-DI in these different cohorts could be attributable to differences in the study designs and participants. Wallace et al (2016) recruited participants from US secondary care and primary care, Khanna et al (2011b) recruited participants who had experienced two or more gout flares in the previous year in a US clinical trial assessing rilonacept vs placebo, whilst both Khanna et al (2011a) and Alvarez-Hernandez et al (2008) included participants from specialist rheumatology clinics. The secondary care setting and inclusion criteria of these existing studies meant that people with more severe gout were recruited compared with this primary care cohort.

7.4.2 Changes observed in disease-specific and generic HRQOL at a group level

The mean changes observed in both disease-specific and generic HRQOL scores at a group level were either small or negligible. This finding is perhaps not surprising considering that this is a study of prevalent cases in primary care, rather than an inception cohort or intervention study.

Since gout flares are the major factor associated with HRQOL (Chandratre et al 2013), the finding that only small or negligible changes were observed in GIS scores at a group level is in keeping with the findings of the previous chapter that the trajectory of gout flares for the majority of participants did not change over time (most participants were allocated to either the infrequent, moderately frequent or frequent and persistent classes).

The small improvement observed at a group level for some HRQOL scores could also be due to change in participants' perceptions of their circumstances due to adaption or adjustment over time (Bowling, 2009; Blome & Augustin, 2015; Fayers & Machin, 2016). HRQOL scores may also change over time because the internal scale that participants use to rate themselves when responding to items has changed (Blome & Augustin, 2015; Fayers & Machin, 2016).

There was negligible change in the SF-36 PF10 scores over time. This is perhaps surprising as lower SF-36 PF10 scores, worse HRQOL, have been associated with older age in the general population (Bowling et al 1999; Brazier et al 1992). However, this is change at a group level and there was more variation in change displayed in the spaghetti plots of SF-36 PF10 scores. It is possible that three years was insufficient time to demonstrate change in physical function in an already older cohort of participants. The findings are consistent with the lack of change demonstrated in the SF-36 physical component scores over two years in participants living with gout receiving 'usual care' in an RCT by Doherty et al (2018).

Worse activity limitation HRQOL was highlighted by HAQ-DI mean change. In this older cohort it is perhaps understandable that HAQ-DI scores would increase for those participants who reported their HAQ-DI scores at both time-points, as HAQ-DI scores have been shown to increase exponentially with age in the general population (Krishnan et al 2004). Worsening disability in older age can be due to a range of factors including the increased burden of comorbidities over time (Manini, 2012).

This chapter compared the mean HRQOL outcomes in the different latent trajectory classes identified in the previous chapter via LCGA. This comparison revealed worse disease-specific and generic HRQOL in the 'frequent and persistent' and 'gradually worsening' classes, in comparison to classes with less frequent flares. This finding is consistent with the findings of previous studies which have reported worse HRQOL with more frequent flares (Becker et al 2009; Chandratre et al 2018; Hirsch et al 2008; Khanna et al 2011a; Khanna et al 2012c; Proudman et al 2019; Scire et al 2013; Stewart et al 2018) and also adds validity to the distinct latent gout flare trajectory classes.

Whilst this chapter has described the change in HRQOL scores over a three-year period, investigation of the various factors potentially associated with change in HRQOL will be addressed in chapter eight.

7.4.3 Individual level change in HRQOL scores

The spaghetti plots displayed change at an individual level and revealed both a wide range of HRQOL scores at baseline and a range of different patterns of change in HRQOL scores over three years. The plotting of intra-individual pattern of change revealed that the GIS WBDA subscale displayed the greatest amount of participants with different scores, in contrast with the MSE subscale which displayed the least amount of different GIS subscale scores. This difference may be, at least in part, attributable to the greater number of items in the GIS WBDA subscale compared to other GIS subscales; thus, permitting a greater number of different scores. Conversely, the GIS MSE has only two items in its scale.

The interpolation lines fitted to the spaghetti plots confirmed the group level pattern of change discussed earlier in this discussion. The interpolation line fitted to the spaghetti plots in this chapter also indicated a linear form for each of the HRQOL scores over time, this finding provides justification for linear modelling of the HRQOL scores.

7.4.4 Correlation and distribution of HRQOL scores

Each of the HRQOL scores were analysed individually to investigate the degree of correlation and dependency between scores at different time-points. The correlation of GIS subscales scores between different time-points was moderate to very high, whilst the SF-36 PF10 and HAQ-DI scores displayed high and very high correlations respectively. This correlation, referred to as auto-correlation or serial correlation, is commonly observed where HRQOL scores are collected on the same participants over time (Fayers & Machin, 2016). It was important to identify this serial correlation within the HRQOL scores in this cohort, as the serial correlation in outcome measures requires special consideration during statistical modelling (Cheng et al 2010; West, Welch & Galecki, 2015) and this investigation confirmed the need to model the HRQOL scores taking the serial correlation in to account in the next chapter.

The investigation of the distribution of the HRQOL scores in this chapter was important in order to inform how they are statistically modelled. Failure to consider non-normally distributed data during modelling can lead to biased parameter estimates and invalid inferences (Dagne & Huang, 2012; Spriensma et al, 2018). The skewed distribution of the SF-36 PF10 and HAQ-DI scores was confirmed by histograms and Q-Q plots, skewed distributions in these outcomes have been reported previously in gout populations (Alvarez-Hernandez et al 2008; Taylor et al 2008). The distribution of GIS subscale scores was closer to a normal distribution compared to the SF-36 PF10 and HAQ-DI scores. Some departure of the outcome from a normal distribution is deemed acceptable in the linear mixed modelling used in the following chapter, as long as the residuals of the model are normally distributed (Cheng et al 2010; West, Welch & Galecki, 2015), thus the distribution of the residuals of linear mixed models used in the next chapter was investigated.

7.4.5 Implications for longitudinal modelling

The results of the exploratory analysis in this chapter influenced the modelling of the GIS subscales, SF-36 PF10 and HAQ-DI scores in the following chapter. The use of linear mixed modelling to investigate factors associated with change in individual GIS subscale scores appeared to be justified based on the identification of a linear form to the GIS scores over time, the correlation between scores from the same GIS subscale, and the lack of skew in the distribution of GIS scores. This investigation also provided justification for modelling each GIS subscale separately, as the analysis displays differences in the magnitude of individual and group level change, correlation, and distribution between different subscales. This justification for modelling the GIS subscales separately is in addition to the fact that the subscales address different conceptual domains relating to the impact of gout.

The linear form and correlation of SF-36 PF10, and HAQ-DI scores also justified the use of linear mixed models to model these outcomes. However, the presence of a skewed

distribution in the SF-36 PF10, and HAQ-DI scores indicated the need to adjust the linear mixed models to take this skewed distribution into account (Cheong, Fotiu & Radenbush, 2001; Maas & Hox, 2004).

7.4.6 Clinical importance of change in HRQOL

The analysis undertaken in this chapter increases our understanding of how HRQOL, in people living with gout, changes over time. Undertaking this investigation in a primary care cohort is of particular relevance to clinical practice as the majority of patients with gout are managed within primary care.

A 'minimal clinically important difference' (MCID) has been defined by Jaeschke, Singer & Guyatt (1989) as the smallest difference in a score on a quality of life domain of interest which patients perceive as beneficial. MCID in people living with gout have been described for the GIS subscales (Khanna et al 2011b), SF-36 PF10 (Khanna et al 2011a) and the HAQ-DI (Singh et al 2016). The MCID in GIS subscale scores ranged from 5 to 8 points (CO, 7.2; UTN, 6.9; WBDA, 5.2; CDA, 7.6) (Khanna et al 2011b). Thus, based on these cut offs only the WBDA GIS subscale changed by a MCID over three years. However, it is important to appreciate that the MCID refers to change at an individual level. Also, Khanna et al (2011b) cautioned against interpreting changes for GIS subscales below the MCID cut offs as not being clinically important, as there is uncertainty relating to the MCID cut offs and the cut offs were derived from RCT data. The MCID cited by Khanna et al (2011a) in SF-36 PF10 scores in people with gout was 5. The SF-36 PF10 scores in this cohort study displayed a mean change which was below the cut off indicating a MCID. The MCID cited by Singh et al (2016) in HAQ-DI scores in people with gout was 0.22. The HAQ-DI scores in this cohort study displayed a mean change which was below the cut off indicating a MCID change in activity limitation. However, these cut offs for MCID in SF-36 PF10 and HAQ-DI scores in people with gout were originally derived in patients with rheumatoid arthritis and scleroderma (Khanna et al 2011a; Singh et al 2016).

In addition, Khanna et al (2011b) determined these MCIDs in a cohort of fewer than 100 people with gout, the majority of whom were prescribed urate-lowering therapy.

Whilst the magnitude of change in HRQOL scores identified in this cohort at a group level may be judged negligible or small, this chapter has shown a wide variation in change in HRQOL scores over time at an individual level. Thus, the levels of change in HRQOL scores seen in this cohort should not be considered clinically irrelevant, simply because the magnitude of change at a group level is negligible or small, or because some HRQOL scores display change at a group level below the levels advocated as a clinically important difference.

7.4.7 Strengths and limitations

This chapter provides a unique insight into change in HRQOL scores in a large cohort of people living with gout in primary care over a three-year period. This analysis of longitudinal HRQOL scores, as opposed to solely cross-sectional analysis, enables the dynamic nature of HRQOL to be captured (Fayers & Machin, 2016; Hahn et al 2007) and addresses the paucity of research investigating change in HRQOL in people living with gout in primary care. The investigation of change in this chapter addressed change in both disease-specific and generic HRQOL measures, thus seeking to capture information relating to a range of HRQOL domains (Hickey et al 2005; Testa & Simonson, 1996). A further strength is that change at both group and individual levels was investigated, using a range of different descriptive analyses. Use of different descriptive analysis techniques, including graphing individual and group level change, has been advocated in order to appreciate the extent of change which can potentially occur in longitudinal data (Leffondre et al 2004; Long, 2012).

It is important to acknowledge that there are limitations in the preliminary exploratory analyses of the longitudinal HRQOL outcomes undertaken.

Some of the analysis methods do not associate participants with their repeated measures. The calculation of measures of central tendency and the creation of box plots, are treating scores at each time-point as being independent. The mean change was calculated by including only participants with scores at two time-points (Locasico & Atri, 2011). Thus, it is important to interpret mean changes with this in mind, as participants who responded at both time-points are likely to be different to the whole cohort overall.

A strength of the analysis undertaken is that SRM and effect sizes were calculated. The calculation of SRM and effect size was justified as they provide information about the magnitude of change in each HRQOL, whilst taking variability in to account and can act as a 'scale free' measure of change thus facilitating comparison of change in scores (Cohen, 1992; Fayers & Machin, 2016; Seidel, Miller & Chow, 2014). The SRM is a specific effect size for use with repeated measures and is widely used for analysis of health outcome questionnaires (Fayers & Machin, 2016; Middel & Van Sonderen, 2002; Morris, 2000; Morris & DeShon, 2002; Wolff Smith & Beretvas, 2009). A challenge when comparing research findings to published effects sizes is that the calculation of effect sizes can be undertaken using a range of different methods, which have the potential to yield different interpretations of change (Middel & Van Sonderen, 2002; Morris & Deshon, 2002; Seidel, Miller & Chow, 2014). When interpreting SRM and effect sizes it is also important to acknowledge that the cut offs proposed by Cohen to interpret effect sizes were derived subjectively and originally intended for use with independent measures (Seidel, Miller & Chow, 2014).

Despite some acknowledged limitations of the analysis in this chapter, it was important to undertake this preliminary analysis, in order to begin to understand the behaviour of HRQOL scores over time. Investigation of the change, functional form, variance, correlation and distribution of the HRQOL scores was important before moving on to analyse the HRQOL scores using more advanced statistical modelling in the following chapter.

7.5 Conclusion

In conclusion, this chapter describes the change in gout-specific HRQOL (GIS subscales) and generic HRQOL (SF-36 PF10, HAQ-DI) scores over a three-year period in a prospective cohort study of people living with gout in primary care. Change in HRQOL scores at a group and individual level was described. At a group level, the magnitude of change in HRQOL scores was small and for some HRQOL scores negligible, however a wide variation in change in HRQOL scores at an individual level was identified. The variation, correlation, and distribution of each HRQOL score over the three-year period was also described. The findings of this chapter have been used to inform the analysis described in the next chapter, where factors associated with change in HRQOL scores will be identified within this cohort using linear mixed models.

8 Chapter Eight Factors associated with change in HRQOL

8.1 Overview of chapter and aim

In the previous chapter, change in gout-specific (GIS subscales) and generic HRQOL (SF-36 PF10, HAQ-DI) scores over a three-year period was described. The aim of the following chapter is to describe the gout-specific, comorbid, socio-demographic and other factors associated with change in disease-specific and generic HRQOL over a three-year period in people living with gout in primary care.

8.2 Background

8.2.1 Factors associated with HRQOL in existing studies

A range of factors have been reported to be associated with disease-specific and/or generic HRQOL in people with gout in previous studies, as described in more detail in section 1.9.4 and summarised in appendix 2. Gout-specific factors associated with HRQOL in previous studies include frequency of flares, experiencing a current or recent flare, the number of joints affected by gout, disease duration, tophi, serum urate, and allopurinol, although studies investigating the relationship between HRQOL and tophi, serum urate levels, and medication to treat gout, have revealed conflicting findings. Other factors associated with HRQOL in people living with gout include the number or type of comorbidities, anxiety, depression, pain, age, sex, level of deprivation, educational attainment, marital status, BMI and alcohol consumption. Gout flares are the major factor associated with impaired disease-specific and generic HRQOL in people living with gout (Becker et al 2009; Chandratre et al 2018; Hirsch et al 2008; Khanna et al 2011a; Khanna et al 2012c; Proudman et al 2019; Scire et al 2013; Stewart et al 2018). The majority of studies are cross-sectional, with a paucity of prospective studies investigating the factors associated with HRQOL particularly in primary care. The previous

studies which have investigated factors associated with change in HRQOL in people with gout have done so over periods no longer than 12 months.

8.3 Method and analysis plan

8.3.1 Aim and Objective

Aim

To describe the factors associated with change in gout-specific and generic HRQOL in people living with gout in primary care over a three-year period.

Objective

1. Use linear mixed modelling to describe the variables associated with change in GIS subscales (CO, MSE, UTN, WBDA, CDA), SF-36 PF10, and HAQ-DI scores over a three-year period.

8.3.2 Data source

The data sources for the analysis presented in this thesis chapter are from the responder and medical record review databases of the prospective cohort study described in chapter three.

8.3.3 Missing data

The investigation of the proportion of missing data in the GIS subscales, SF-36 PF10 and HAQ-DI and also covariates was undertaken in chapter four. LMMs tolerate a degree of missing data, as they allow participants to have an unequal number of measurements and the measurements to have been collected at unevenly spaced time-points (Cnaan, Laird & Slasor, 1997; Long, 2012; West, Welsh & Galecki, 2015). Thus, participants were included within the LMM analysis if they had an outcome response on a minimum of one time-point (Long, 2012).

8.3.4 Selection of potential covariates

Covariates were selected based on whether they have been associated with HRQOL in people with gout in previous studies (including this cohort's baseline analysis paper) and were introduced in chapter one. This consideration of existing research when selecting predictors in linear mixed modelling has been advocated as good practice (Cheng et al 2010). Table 8.1 lists the range of variables which were included as covariates. These covariates included gout-specific, comorbid, socio-demographic and other factors.

Continuous variables were added to the LMM as covariates in their continuous form, aiming to maximise the predictive information provided by that variable (Royston et al 2009).

For non-time varying covariates, an interaction between each covariate and time was added to the model. This addition of an interaction between time and covariates was used to test whether the impact of any of the covariates on the outcomes changed over time (Berrington, Smith & Sturgis, 2006).

Table 8.1 Selection of covariates in linear mixed models for GIS subscales, SF-36 PF10 and HAQ-DI

	Time-varying covariate?	**Variable format**	**Interpretation of reference group**
Gout-specific			
Number of gout flares[†]	✓	Categorical	No gout flares
Current flare	✓	Dichotomous	No current flare
Oligo/polyarticular flares	✓	Dichotomous	No history of oligo/polyarticular flares
Disease duration	×	Continuous	-
Allopurinol use (self-reported)	✓	Dichotomous	No allopurinol use
Comorbidities			
Number of self-reported comorbidities	×	Continuous	-
eGFR <60 mL/min/1.73m^2 [◇]	×	Dichotomous	eGFR ≥ 60
0-10 NRS pain over previous week	✓	Continuous	-
Presence of body pain	×	Dichotomous	No body pain
PHQ-9 score	×	Continuous	-
GAQ-7 score	×	Continuous	-
Socio-demographic			
Age	×	Continuous	-
Sex	×	Dichotomous	Male
Ethnicity	×	Dichotomous	Caucasian
Indices of multiple deprivation (IMD) decile	×	Continuous	-
Married or cohabiting	×	Dichotomous	Not married or cohabiting
Attendance at further education	×	Dichotomous	No attendance at further education
BMI and alcohol			
BMI	✓	Continuous	-
Alcohol consumption	×	Categorical	Alcohol consumption-Never

✓**Time-varying** covariate in questionnaire at baseline, 6, 12, 24 and 36 months

[†] **Number of gout flares** in previous 12 months at baseline, 12 months and 36 months, and in previous 6 months at 6 and 12 months ◇From **medical records** in the two years prior to baseline

Total number of self-reported comorbidities are number of comorbidities, from pre-specified list, self-reported in baseline questionnaire

eGFR <60 mL/min/1.73m^2 indicative of chronic kidney disease stage ≥ 3

NRS pain in last week ranges from 0 (no pain) to 10 (pain as bad as it can be)

Body pain (including ache or discomfort or stiffness) for one day or longer in the 4 weeks prior to baseline

PHQ-9 score ranges from 0 to 27 higher score indicates more severe depression

GAD-7 score ranges from 0 to 21 higher score indicates more severe anxiety

BMI body mass index kg/m^2

IMD Indices of multiple deprivation highest score indicates least deprived area

8.3.5 Linear Mixed Modelling

A linear mixed model (LMM) is a linear statistical model which can be applied to parametric clustered, repeated measures, and longitudinal data (West, Welch & Galecki, 2015). Such models are extensions of traditional regression models (Long, 2012) and aim to quantify the relationship between a continuous outcome variable and covariates (West, Welch & Galecki, 2015). The models are referred to as 'mixed' as they include a mixture of both fixed and random effects (Locascio & Atri, 2011; West, Welch & Galecki, 2015). The LMMs used in this chapter can also be referred to as multi-level or hierarchical models as the outcome measures are clustered within participants (Twisk, 2006). Mixed models specifically take this hierarchy and dependency of observations into account when analysing longitudinal data (Twisk, 2006). Alternative names for linear mixed models used in the literature include multilevel models, random effects models, hierarchical models, mixed effects models (Kirkwood & Sterne, 2003; Twisk, 2006), linear mixed effects regression (Long, 2012), linear mixed effects models (Muth et al 2016), and multivariate/multivariable mixed effects regression (Singh et al 2016; Wallace et al 2016).

The decision to analyse the GIS subscales, SF-36 PF10 and HAQ-DI scores using LMM was informed by the analysis in previous chapters. LMM was used in this analysis because it allows missing data (Cheng et al 2010; Long, 2012; West, Welch & Galecki, 2015), auto-correlation/ serial correlation (Cheng et al 2010; West, Welch & Galecki, 2015), different covariance structures (Cheng et al 2010; Muth et al 2016), the inclusion of time-varying covariates (Long, 2012; West, Welch & Galecki, 2015), and data which are collected at irregularly spaced intervals (Cheng et al 2010; Locascio & Atri, 2011; Long, 2012).

Fixed effects

The fixed effects components in LMMs are called regression coefficients (represented by the symbol β) and are similar to regression coefficients in traditional regression (Cheng et al 2010;

West, Welch & Galecki, 2015). These fixed effects regression coefficients provide estimates of the average response in the outcome variable for the whole group, and consequently are population specific estimates (Cheng et al 2010; Long, 2012; West, Welch & Galecki, 2015). Thus, the fixed effects component of a LMM describes the relationship of a covariate to the whole population (West, Welch & Galecki, 2015). As fixed effects describe the mean response for the whole population, they are often the primary interest of clinical and epidemiological research (Cheng et al 2010).

Random effects

In a LMM with random effects, conceptually each participant will have their own intercept and slope describing the relationship of interest, however in practice the regression coefficients for the intercept and the slope for each participant are not individually estimated; because this would be inefficient, as it would generate a large number of parameters (Ghisletta et al 2015; Twisk, 2006). Instead, in a LMM with random effects (a random intercept and slope), the variance of intercepts, the variance of the slopes and the covariance of these parameters are estimated (Ghisletta et al 2015; Twisk, 2006).

Residual random error

Random error in the model represents the error in the measurement of the outcome at an individual level (Long, 2012; West, Welch & Galecki, 2015). Thus, the error variance estimated for all subjects reflects the within-subject error variance (Long, 2012).

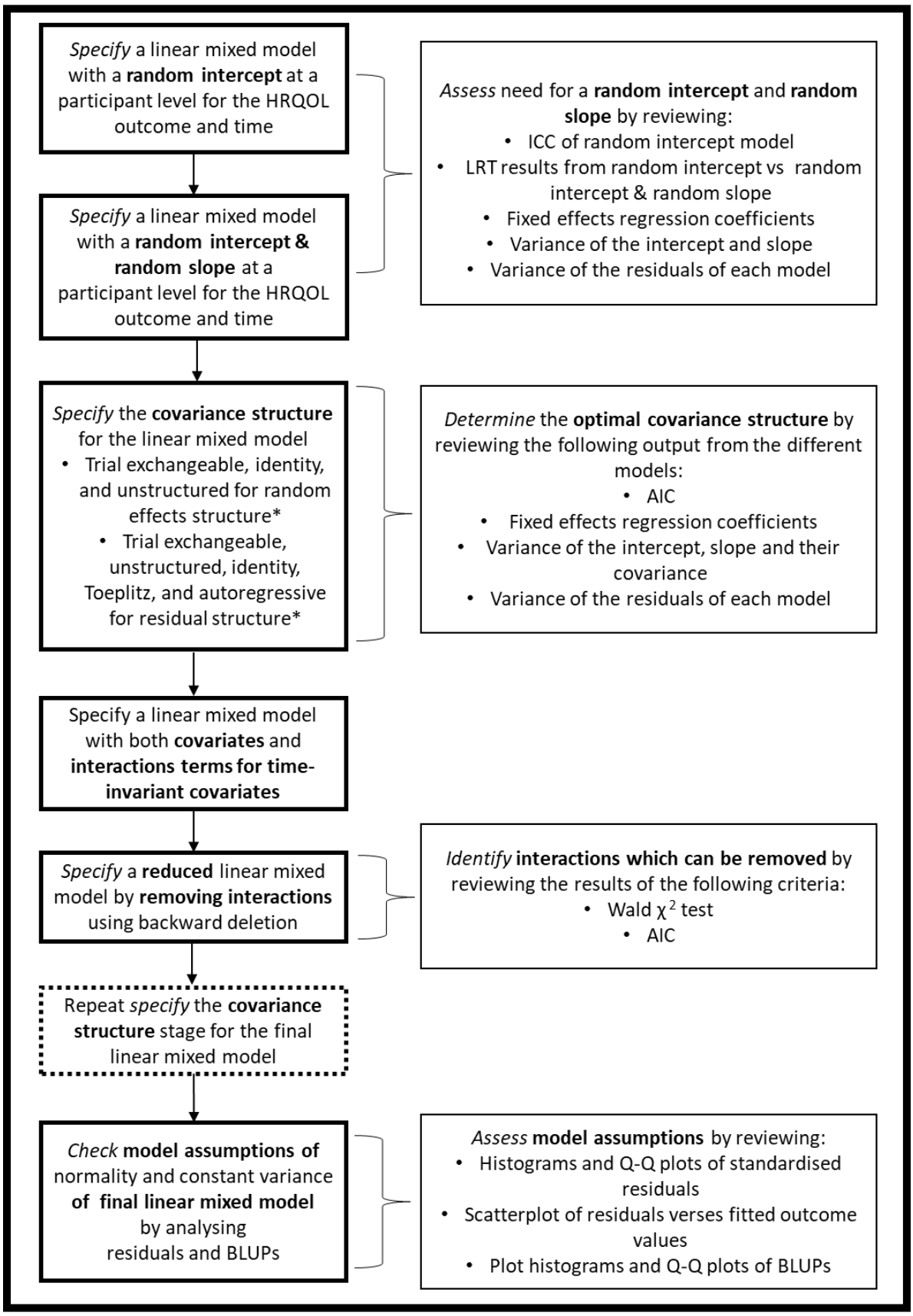
Stages of linear mixed modelling used in this chapter

The stages of linear mixed modelling undertaken in this chapter can be found in Figure 8.1. The GIS CO, GIS MSE, GIS UTN, GIS WBDA, GIS CDA, SF-36 PF10 and HAQ-DI were each analysed in separate linear mixed models.

The linear mixed modelling was undertaken using STATA version 14. Table 8.2 describes the range of criteria which were considered during the linear mixed modelling process, due to the absence of one widely accepted model fit criterion (Gurka, 2006; Wang & Schaalje, 2009).

Estimation method and robust standard errors

Residual maximum likelihood estimation (REML) is a method which is often used when estimating parameters in LMM, due to its ability to produce unbiased estimates of the covariance parameters (Gurka, 2006; West, Welsh & Galecki, 2015). Thus, REML was used when modelling GIS outcomes. However, the analysis in chapter six revealed that the SF-36 PF10 and HAQ-DI scores displayed a left and right skew respectively. If data are not normally distributed, estimates and tests will not be robust to the violation of distribution assumptions (Cheong, Fotiu & Radenbush, 2001). Thus, the LMMs for the SF-36 PF10 and HAQ-DI scores were amended to use robust standard errors, in order to deal with the non-normal distribution of these outcomes (Cheong, Fotiu & Radenbush, 2001; Maas & Hox, 2004) and maximum likelihood estimation (MLE), rather than REML, was used to enable robust standard errors to be used. Whilst REML and MLE have different strengths and weaknesses as estimation methods, they are likely to yield similar results in large samples (Rabe-Hesketh & Skrondal, 2008; Twisk, 2006).



*** descriptions of the different covariance structures** trialled can be found in appendix 15

AIC, Akaike Information Criteria; **BLUPs**, best linear unbiased predictors; **ICC**, Intraclass correlation coefficient; **LRT**, Likelihood Ratio Test; **Q-Q plot**, quantile-quantile plot; **Wald χ^2**, Wald Chi-Squared Statistic.

Figure 8.1 Stages of linear mixed modelling undertaken in this chapter

Table 8.2 Justification for criteria used in linear mixed modelling and guidance for interpretation of results

Criteria	Justification and guidance
Intraclass correlation coefficient (ICC)	A significant ICC suggested that there were differences in the dependent variable between individuals (Garson, 2013), providing justification for a random intercept, whereas an ICC closer to zero suggested a lack of difference between individuals (Garson, 2013).
Likelihood Ratio Test (LRT)	A statistically significant LRT result ($p < 0.05$) indicated an improvement in model fit with the inclusion of a random slope compared to a nested model with just a random intercept (West, Welch & Galecki, 2015). However, caution should be taken when interpreting LRT in linear mixed modelling as LRT results can be conservative with a tendency of P values to be large, thus leading to the selection of a model with less random effects (Long, 2012; West, Welch & Galecki, 2015).
Fixed effects **regression coefficients**	The fixed effects regression coefficients (95% CI) were reviewed to note any change when i) comparing models with and without random slopes, and ii) when comparing models with different covariance structures.
Variance of intercept, slope and **covariance**	The variance of the intercept, variance of slope and covariance of intercept with slope (along with 95 % CI) were reviewed to note any change (Twisk, 2006) when i) comparing models with and without random slopes, and ii) when comparing models with different covariance structures.
Variance of residuals	The variance of the residuals of each model were reviewed to see how changing the model changed the magnitude of the variance of the residuals (Twisk, 2006).
Information criteria i.e. AIC	Information criteria, Akaike Information Criteria (AIC) (Akaike, 1987), was used to select the covariance structure used in the models (Garson, 2013; Wang & Schaalje, 2009). Thus, the AIC results for models with different covariance structures for random effects and residuals were compared; with a smaller AIC result indicating a better fitting model (Garson, 2013; Gurka, 2006). AIC (Akaike, 1987) was also used to identify covariate interactions which could be removed from the model via backward deletion. The AIC for models with and without particular interactions were compared, with a smaller result indicating a better fitting model (Garson, 2013; Gurka, 2006).
Wald χ^2 test	Wald χ^2 test can be used when the full model has been fitted to identify fixed effects which could be removed from the linear mixed model (West, Welch & Galecki, 2015). Thus, the Wald test, along with AIC, was used to select interactions which could be removed from the model. A statistically significant result ($p < 0.05$) indicated that a fixed effect (interaction) could be retained in the model (West, Welch & Galecki, 2015). As REML was used in the LMM the Wald χ^2 test was used as an alternative to LRT (StataCorp, 2013).
Standardised **residuals**	Standardised residuals were assessed to check the assumptions of the final LMMs (Cheng et al 2010; West, Welch & Galecki, 2015). Histograms and Q-Q plots of standardised residuals were created and scrutinised to investigate whether the residuals were normally distributed (Cheng et al 2010). A scatterplot of residuals verses fitted outcome values were created and scrutinised to investigate whether the residuals had constant variance.
Best linear unbiased predictors (**BLUPs)**	BLUPs, the predicted values of the random effects given the observed data, were plotted in histograms and Q-Q plots to investigate normality of the random effects (West, Welch & Galecki, 2015).

8.4 Results

8.4.1 Specifying the linear mixed models

Random intercept linear mixed models

The LMM with random intercepts for each GIS subscale (Table 8.3), the SF-36 PF10 and HAQ-DI (Table 8.5) demonstrated variance at an intercept level across all participants. The ICC results ranged from 0.48 to 0.75 for the GIS subscales (Table 8.3) and were 0.73 for the SF-36 PF10 and 0.86 for the HAQ-DI (Table 8.5), suggesting there were differences in the HRQOL scores between individuals. These findings justified the retention of a random intercept in the LMMs.

Random intercept and slope linear mixed models

The LRTs comparing the LMM models with both a random intercept and random slope, to a model with only a random intercept were statistically significant for each of the GIS subscales (GIS CO, MSE, WBDA p <0.001; GIS UTN p 0.03; GIS CDA p 0.003) (Table 8.3), the HAQ-DI (p <0.001) but not the SF-36 PF10 (Table 8.5). The statistically significant LRT results for the GIS subscales and the HAQ-DI indicated an improvement in model fit with the inclusion of more random effects (i.e. random slope in addition to random intercept) compared to a nested model with less random effects (i.e. random intercept only model). The LMM with both a random intercept and random slope for each GIS subscale demonstrated variance at both an intercept and slope level (Table 8.3). The addition of a random slope reduced the residual error in each GIS subscale LMM (Table 8.3) in comparison with the residual error in the GIS LMM with a random intercept only (Table 8.3). Although the variance at a slope level for the HAQ-DI LMM was small, the confidence interval did not cross zero (Table 8.5). The HAQ-DI model with both a random intercept and random slope also returned a slightly lower AIC, than the random intercept only model, indicating a better fitting model (Table 8.5).

Table 8.3 Comparison of results from GIS subscales LMMs with random intercept only, and random intercept with random slope independent covariance

	GIS CO n=1167	**GIS MSE** n=1156	**GIS UTN** n=1143	**GIS WBDA** n=1166	**GIS CDA** n=1161
Random intercept independent covariance					
Intercept β (95% CI)	**49.13 (47.61, 50.65)**	**40.68 (39.26, 42.10)**	**33.47 (32.41, 34.53)**	**44.39 42.95, 45.83**	**40.31 (39.00, 41.63)**
Slope β (95% CI)	**-0.17 (-0.21, -0.12)**	**-0.07 (-0.12, -0.03)**	**-0.07 (-0.12, -0.03)**	**-0.16 (-0.20, -0.13)**	-0.03 (-0.07, 0.01)
Variance intercept (95% CI)	**544.31 (494.42, 599.24)**	**446.83 (404.26, 493.89)**	**194.52 (172.82, 218.94)**	**518.16 (472.29, 568.49)**	**406.42 (368.97, 447.67)**
Variance residuals (95% CI)	**241.11 (228.70, 254.19)**	**251.11 (238.11, 264.81)**	**212.62 (201.49, 224.35)**	**169.09 (160.33, 178.33)**	**181.72 (172.34, 191.60)**
Intraclass correlation coefficient (ICC) (95% CI)	**0.69 (0.67, 0.72)**	**0.64 (0.61, 0.67)**	**0.48 (0.44, 0.51)**	**0.75 (0.73, 0.77)**	**0.69 (0.67, 0.72)**
Akaike Information Criteria (AIC)	34957.76	34445.34	32647.10	33592.94	33727.52
Random intercept and slope independent covariance					
Intercept β(95% CI)	**49.13 (47.62, 50.63)**	**40.67 (39.27, 42.08)**	**33.47 (32.42, 34.52)**	**44.39 (42.96, 45.82)**	**40.32 (39.01,41.63)**
Slope β(95% CI)	**-0.17 (-0.21, -0.12)**	**-0.07 (-0.12, -0.03)**	**-0.07 (-0.12, -0.03)**	**-0.16 (-0.20, -0.12)**	-0.03 (-0.07, 0.01)
Variance intercept (95% CI)	**545.78 (495.29, 601.42)**	**444.71 (401.67, 492.36)**	**194.31 (172.35, 219.07)**	**518.62 (472.39, 569.34)**	**405.14 (367.51, 446.61)**
Variance slope (95% CI)	**0.10 (0.06, 0.15)**	**0.11 (0.07, 0.16)**	**0.03 (0.01, 0.08)**	**0.07 (0.04, 0.105)**	**0.04 (0.02, 0.08)**
Variance residuals (95% CI)	**221.17 (208.05, 235.10)**	**230.09 (216.39, 244.65)**	**206.06 (193.87, 219.01)**	**155.81 (146.46, 165.76)**	**173.65 (163.42, 184.52)**
Likelihood ratio test (LRT)	**p <0.001**	**p<0.001**	**p 0.03**	**p<0.001**	**p0.003**
Akaike Information Criteria (AIC)	34936.01	34421.93	32645.97	33576.04	33725.87

Table 8.4 GIS subscales LMMs with random intercept and random slope unstructured covariance results

	GIS CO n=1167	**GIS MSE** n=1156	**GIS UTN** n=1143	**GIS WBDA** n=1166	**GIS CDA** n=1161
Random intercept and slope unstructured covariance					
Intercept β (95% CI)	**49.13 (47.59, 50.67)**	**40.68 (39.24, 42.12)**	**33.54 (32.44, 34.64)**	**44.39 42.94, 45.85**	**40.33 (39.00, 41.66)**
Slope β (95% CI)	**-0.17 (-0.21, -0.12)**	**-0.07 (-0.12, -0.03)**	**-0.08 (-0.12, -0.04)**	**-0.16 (-0.20, -0.12)**	-0.03 (-0.07, 0.01)
Variance intercept (95% CI)	**581.03 (524.64, 643.48**	**474.01 (425.09, 528.56)**	**233.52 (204.72, 266.38)**	**538.43 (488.41, 593.58)**	**421.05 (379.16, 467.56)**
Variance Slope (95% CI)	**0.13 (0.09, 0.19)**	**0.14 (0.10, 0.20)**	**0.09 (0.06, 0.15)**	**0.08 (0.05, 0.12)**	**0.06 (0.03, 0.10)**
Covariance Slope, Intercept (95%CI)	**-2.14 (-3.48, -0.79)**	**-1.73 (-3.03, -0.44)**	**-2.08 (-3.03, -1.12)**	**-1.19 (-2.23, -0.15)**	-0.85 (-1.81, 0.11)
Variance residuals (95% CI)	**216.09 (203.20, 229.79)**	**224.75 (211.13, 239.25)**	**195.26 (183.36, 207.94)**	**153.47 (144.19, 163.35)**	**171.18 (160.89, 182.14)**
Akaike Information Criteria (AIC)	34927.48	34416.44	32626.44	33572.74	33724.72

Table 8.5 Comparison of results for both SF-36 PF10 and HAQ-DI LMMs with random intercept only and random intercept with random slope

	SF-36 PF10 n=1097	**HAQ-DI** n=1168	**HAQ-DI** n=1168
Random intercept	independent covariance		
Intercept β (95% CI)	**74.73 (73.22, 76.24)**	**0.51 (0.47, 0.55)**	-
Slope β (95% CI)	0.01 (-0.03, 0.05)	**0.00 (0.00, 0.00)**	-
Variance intercept (95% CI)	**516.21 (471.37, 565.32)**	**0.45 (0 .40, 0.50)**	-
Variance residuals (95% CI)	**192.73 (159.04, 233.55)**	**0.07 (0.06, 0.09)**	-
Intraclass correlation coefficient (ICC) (95% CI) (calculated in model without robust standard errors)	**0.73 (0.70, 0.752)**	**0.86 (0.85, 0.87)**	-
AIC	28481.42	4324.949	-
Random intercept & random slope	independent covariance		unstructured covariance
Intercept β (95% CI)	**74.73 (73.22, 76.24)**	**0.51 (0 .47, 0.55)**	**0.51 (0.47, 0.55)**
Slope β (95% CI)	0.01 (-0.03, 0.05)	**0.00 (0 .00, 0.00)**	**0.00 (0 .00, 0.00)**
Variance intercept (95% CI)	**516.21 (471.37, 565.31)**	**0.45 (0 .40, 0.50)**	**0.45 (0.40, 0.50)**
Variance slope (95% CI)	0.00 (0.00, 1.63)	**0.00 (0.00, 0.00)**	**0.00 (0.00, 0.00)**
Covariance slope, intercept (95%CI)	-	-	0.00 (-0.00, 0.00)
Variance residuals (95% CI)	**192.73 (159.04, 233.55)**	**0.07 (0.06, 0.08)**	**0.07 (0.06, 0.08)**
Likelihood ratio test (LRT)	P=1.00	**P<0.001**	-
AIC	28483.42	4312.898	4314.24

Unstructured covariance was not explored for the SF-36 PF10 as the best fitting model for the SF-36 PF10 was the intercept only model.

The residual error for the HAQ-DI LLM model was the same after the addition of a random slope, however the confidence interval did narrow slightly (Table 8.5). There was no change in the residual error in the SF-36 PF10 LMM with the addition of a random slope (Table 8.5). There was little difference in the AIC for the SF-36 PF10 models with and without a random slope but the AIC for the random intercept only model was 1.68 lower indicating a better fitting model (Table 8.5). These findings justified the retention of a random slope, in addition to a random intercept, in the LMMs for the GIS subscales and the HAQ-DI but justified the use of a random intercept only LMM for SF-36 PF10 scores.

Covariance structure

The AIC results for each of the GIS LMMs with an unstructured covariance were lower (Table 8.4) indicating a better fitting model, than the AIC results for the GIS LMMs with independent covariance structures (Table 8.3). The variance of the residuals for all GIS LMMs reduced by small amounts, after the addition of the unstructured covariance (Table 8.4). The findings justified the retention of an unstructured covariance for the random effects. A negative covariance was identified for GIS CO, GIS MSE, GIS UTN, and GIS WBDA LMMs (Table 8.4) indicating that the higher a person's intercept the smaller the rate of change (Twisk, 2006).

The addition of an unstructured covariance did not change the residual error in the HAQ-DI LMM. The confidence intervals for the covariance estimated in the HAQ-DI LMM with an unstructured covariance also crossed zero. The AIC for the HAQ-DI LMM with an independent covariance structure was slightly lower than the AIC for the HAQ-DI LMM with an unstructured covariance. These results suggested that an unstructured covariance may not be necessary, however the HAQ-DI LMM model with an independent covariance plus the addition of covariates and all interactions failed to run/ converge (repeatedly displaying the warning 'back up'). Therefore, the unstructured covariance was used to reduce the model and the covariance structure was then reassessed after reducing the model.

For both the GIS subscale and HAQ-DI models the other covariance structures (exchangeable, identity) trialled for the random effects (slope and intercept covariance) returned higher AIC results and greater residual variance. Also covariance structures for the residuals trialled were not retained in the model as either the AIC or variance of the residual did not improve (autoregressive covariance) or the model did not run (exchangeable, unstructured, identity, Toeplitz).

Reducing the model

The interactions for the following time-invariant covariates and time were retained in the GIS LMM models due to a Wald test result of <p 0.05; sex x time (GIS CO model), alcohol x time (GIS UTN), non-Caucasian x time (GIS CDA), GAD-7 score x time (GIS CDA) and indices of multiple deprivation x time (both SF-36 PF10 and HAQ-DI models).

Checking the covariance structure of the full model

After reducing the models, the unstructured covariance was retained for the GIS subscale and HAQ-DI LMMs, as this covariance yielded lower residual error variance in comparison to an independent covariance structure. Also, a lower AIC was observed for the GIS CO, GIS UTN and GIS WBDA when an unstructured covariance structure was used compared to an independent structure and there was little difference in the AIC observed for the GIS MSE and GIS CDA models with an independent or unstructured covariance structure. There was also little difference in the AIC observed for the HAQ-DI LMM with an independent covariance structure (just 0.8 lower), compared to the unstructured covariance.

Other covariance structures trialled either returned a greater AIC or variance of the residuals or the LMM did not run.

8.4.2 Factors associated with change in disease-specific HRQOL

GIS Concern overall

The following gout-specific factors were associated with a deterioration in HRQOL over three years, measured using the GIS Concern overall (GIS CO) subscale: frequency of gout flares (this was dose-dependent with stronger associations observed with more flares), experiencing a current flare, and a history of oligo/polyarticular flares (Table 8.6). Longer disease duration at baseline was associated with an improvement in HRQOL. No association was observed between allopurinol use and change in HRQOL.

Greater pain, the presence of body pain, and worse anxiety were associated with a deterioration in HRQOL, measured via the GIS CO, over three years (Table 8.6). However, no associations with the number of comorbidities, eGFR <60, or depression severity, and change in HRQOL were observed.

As shown in Table 8.6 older age was associated with an improvement in HRQOL, measured via GIS CO, over three years. There was an association between the interaction between sex with time and HRQOL indicating that the relationship between sex and HRQOL changed over time. Change in GIS CO over three years was not associated with being non-Caucasian, BMI, IMD, married or cohabiting, attendance at further education, and alcohol consumption.

When adjusted for gout-specific, comorbid, socio-demographic factors, BMI and alcohol frequency, time was associated with an improvement in HRQOL measured via GIS CO (see Table 8.6).

Table 8.6 Factors associated with change in disease-specific HRQOL; results from GIS subscale LMMs with random intercept and random slopes with unstructured covariance

	GIS CO n=595 (β (95% CI))	**GIS MSE** n=595 (β (95% CI))	**GIS UTN** n=591 (β (95% CI))	**GIS WBDA** n=593 (β (95% CI))	**GIS CDA** n=595 (β (95% CI))
Intercept	**68.14(53.82, 82.47)**	**43.23(27.84, 58.64)**	**32.13(21.48, 42.78)**	**65.27(50.45, 80.09)**	**38.81(25.35, 52.28)**
Time	**-0.14(-0.20, -0.08)**	-0.06(-0.12, 0.00)	0.10(-0.10, 0.30)	**-0.17(-0.23, -0.12)**	0.03(-0.03, 0.09)
Gout-specific					
0 gout flares†◊	reference	reference	Reference	reference	reference
1 gout flare†◊	**6.29(4.04, 8.54)**	**3.36(0.98, 5.74)**	**4.42(2.33, 6.52)**	-1.69(-3.71, 0.33)	**2.60(0.56, 4.64)**
2 gout flares†◊	**11.69(8.94, 14.43)**	**4.94(2.01, 7.86)**	**7.94(5.43, 10.46)**	-1.36(-3.84, 1.11)	**4.81(2.32, 7.31)**
3 gout flares†◊	**12.09(8.67, 15.50)**	**7.26(3.68, 10.84)**	**8.52(5.39, 11.66)**	0.59(-2.50, 3.68)	**3.27(0.17, 6.37)**
4 gout flares†◊	**16.30(12.01, 20.59)**	**7.85(3.36, 12.34)**	**11.80(7.87, 15.72)**	1.46(-2.39, 5.30)	**4.67(0.78, 8.55)**
≥5 gout flares†◊	**19.31 (15.65, 22.97)**	**10.14(6.29, 14.00**)	**12.17(8.88, 15.45)**	1.28(-2.07, 4.64)	**7.87(4.54, 11.20)**
Current flare◊	**4.24(1.12, 7.35)**	-0.34(-3.64, 2.95)	**7.01(4.08, 9.94)**	**-3.30(-6.06, -0.54)**	-0.45(-3.27, 2.36)
Oligo/polyarticular flares◊	**4.88(2.77, 6.99)**	**2.86(0.62, 5.09)**	-0.22(-2.06, 1.61)	**3.46(1.52, 5.41)**	**3.34(1.42, 5.26)**
Allopurinol use◊	-0.84(-3.52, 1.85)	1.34(-1.52, 4.20)	**-7.13(-9.28, -4.99)**	**2.63(0.03, 5.24)**	2.35(-0.13, 4.83)
Disease duration	**-0.27(-0.41, -0.13)**	**-0.17(-0.32, -0.01)**	**-0.12(-0.22, -0.02)**	-0.11(-0.26, 0.04)	**-0.16(-0.29, -0.02)**
Comorbidities					
Total number of Comorbidities	-0.08(-1.33, 1.17)	-0.11(-1.45, 1.24)	0.59(-0.30, 1.47)	-0.09(-1.41, 1.23)	0.45(-0.73, 1.63)
eGFR <60mL/min/1.73m^2	1.34(-2.62,5.31)	0.52(-3.77, 4.80)	-2.02(-4.77, 0.73)	1.47(-2.74, 5.70)	2.16(-1.59, 5.90)
Pain NRS score◊	**1.24(0.82, 1.66)**	0.43(-0.02, 0.87)	**0.64(0.26, 1.01)**	**1.61(1.22, 1.99)**	**0.54(0.16, 0.92)**
Body pain	**5.23(1.74, 8.73)**	**7.69(3.91, 11.47)**	1.92(-0.50, 4.35)	2.59(-1.11, 6.29)	1.97(-1.33, 5.27)
PHQ-9 score	0.50(-0.02,1.01)	**0.61(0.06, 1.17)**	0.32(-0.04, 0.68)	**1.10(0.55, 1.64)**	**0.82(0.34, 1.31)**
GAD-7 score	**0.61(0.03, 1.19)**	**0.72(0.10, 1.35)**	-0.04(-0.44, 0.36)	0.40(-0.21, 1.01)	**1.22(0.66, 1.77)**
GAD-7 X Time	-	-	-	-	**-0.01(-0.03, -0.00)**

β linear mixed model coefficient **†** in previous 12 months at baseline, 12 months and 36 months, in previous 6 months at 6 and 12 months **◊Time-varying** covariate in questionnaire at baseline, 6, 12, 24 and 36 months. **Number of comorbidities** total number of self-reported comorbidities in questionnaire; **eGFR** <60 mL/min/1.73m^2 indicative of chronic kidney disease; **NRS pain in last week** ranges from 0 (no pain) to 10 (pain as bad as it can be)**; Body pain** (including ache or discomfort or stiffness) for one day or longer in the 4 weeks prior to baseline; **PHQ-9 depression score** ranges from 0 (minimal) to 27 (severe); **GAD-7 anxiety score** ranges from 0 (minimal) to 21 (severe).

Each **GIS subscale** scored from **0 to 100;** higher scores on each scale indicating a greater impact of gout on HRQOL/worse HRQOL, **positive β** coefficient = deterioration, and **negative β** coefficient = improvement.

Table 8.6 cont. Factors associated with change in disease-specific HRQOL; results from GIS subscale LMMs with random intercept and random slopes with unstructured covariance

	GIS CO n=595 (β (95% CI))	**GIS MSE** n=595 (β (95% CI))	**GIS UTN** n=591 (β (95% CI))	**GIS WBDA** n=593 (β (95% CI))	**GIS CDA** n=595 (β (95% CI))
Socio-demographic					
Age	**-0.53(-0.69, -0.37)**	**-0.27(-0.45, -0.10)**	-0.03(-0.15, 0.08)	**-0.44(-0.61, -0.27)**	**-0.26(-0.41, -0.10)**
Female	0.73(-4.66, 6.13)	-1.03(-6.46, 4.40)	2.54(-1.00, 6.07)	1.10(-4.24, 6.44)	2.03(-2.72, 6.77)
Female x Time	**0.19(0.00, 0.37)**	-	-	-	-
Non-Caucasian	1.39(-10.37, 13.16)	4.37(-8.23, 16.97)	5.02(-3.31, 13.35)	1.00(-11.35, 13.36)	1.09(-10.58, 12.75)
Non-Caucasian X Time	-	-	-	-	**0.75(0.27, 1.22)**
IMD	0.25(-0.31, 0.81)	0.37(-0.23, 0.98)	0.28(-0.11, 0.67)	0.36(-0.24, 0.95)	0.19 -0.34, 0.72)
Married or cohabiting	2.47(-1.41,6.35)	0.01(-4.20, 4.18)	1.04(-1.67, 3.75)	-0.11(-4.22, 4.01)	-1.52(-5.18, 2.14)
Further education	-1.15(-4.80, 2.51)	-1.56(-5.53, 2.40)	**-2.58(-5.10, -0.06)**	-3.87(-7.76, 0.03)	-3.42(-6.88, 0.04)
BMI & alcohol					
BMI◊	-0.03(-0.25, 0.20)	-0.02(-0.26, 0.21)	-0.12(-0.30, 0.06)	0.03(-0.18, 0.24)	0.20(-0.01, 0.40)
Alcohol - Never	reference	reference	Reference	reference	reference
Special occasions	-5.18(-12.45, 2.08)	-3.85(-11.43, 3.74)	2.67(-3.18, 8.52)	-6.89(-14.30, 0.52)	-5.27(-11.89, 1.35)
1-3 times/month	-5.18(-12.43, 2.08)	1.32(-6.52, 9.17)	2.33(-3.66, 8.32)	-7.32(-15.02, 0.39)	-6.10(-12.95, 0.75)
1-2 times/week	-1.03(-7.64, 5.58)	0.57(-6.57, 7.71)	-0.28(-5.40, 5.15)	-2.30(-9.29, 4.69)	-2.33(-8.56, 3.91)
3-4 times/week	-1.25(-7.82, 5.32)	1.96(-5.15, 9.07)	-2.10(-7.48, 3.27)	-5.97(-12.91, 0.98)	-3.31(-9.51, 2.89)
Daily or almost daily	-4.73(-11.13, 1.66)	-3.30(-10.22, 3.62)	-0.02(-5.48, 5.04)	**-7.53(-14.30, -0.75)**	-4.8(-10.92, 1.16)
Alcohol - Never	reference	reference	Reference	reference	reference
Special occasions x Time	-	-	-0.15(-0.41, 0.11)	-	-
1-3 times/month x Time	-	-	-0.20(-0.45, 0.06)	-	-
1-2 times/week x Time	-	-	-0.10(-0.33, 0.14)	-	-
3-4 times/week x Time	-	-	-0.13(-0.35, 0.09)	-	-
Daily x Time	-	-	**-0.23(-0.46, -0.01)**	-	-

β linear mixed model coefficient ◊**Time-varying covariate** in questionnaire at baseline, 6, 12, 24 and 36 months.

BMI body mass index kg/m^2; **IMD** Indices of multiple deprivation, highest score indicates least deprived.

Each **GIS subscale** scored from **0 to 100;** higher scores on each scale indicating a greater impact of gout on HRQOL/ worse HRQOL, **positive β** coefficient = deterioration, and **negative β** coefficient = improvement.

GIS Medication side effects

Over three years a deterioration in HRQOL, measured via GIS Medication side effects (GIS MSE) subscale, was associated with the frequency of gout flares (this was dose-dependent with stronger associations observed with more flares), and a history of oligo/polyarticular flares (Table 8.6). However, no association was observed between change in HRQOL and experiencing a current flare, or using allopurinol. Longer disease duration at baseline was associated with an improvement in HRQOL.

The presence of body pain, worse depression, and worse anxiety were associated with a deterioration in HRQOL, measured via the GIS MSE (Table 8.6). Change in GIS MSE over three years was not associated with the number of comorbidities, eGFR <60, or greater pain.

Older age was associated with an improvement in HRQOL over three years (Table 8.6). No association between change in HRQOL and sex, being non-Caucasian, BMI, IMD, married or cohabiting, attendance at further education, and alcohol consumption was observed.

After adjustment for gout-specific, comorbid, socio-demographic factors, BMI and alcohol frequency, time was not associated with an improvement in HRQOL measured by GIS MSE (Table 8.6).

GIS Unmet treatment need

A deterioration in HRQOL, measured via the GIS Unmet treatment need (GIS UTN) subscale, over three years in this cohort was associated with the frequency of gout flares (this was dose-dependent with stronger associations observed with more flares), and also experiencing a current flare (Table 8.6). Longer disease duration at baseline, and allopurinol use were associated with an improvement in HRQOL. No association was observed between a history of oligo/polyarticular flares and change in HRQOL.

No association in change in HRQOL, measured via GIS UTN, and the number of comorbidities,

eGFR <60, the presence of body pain, depression severity or anxiety severity were observed (Table 8.6). However, greater pain was associated with a deterioration in GIS UTN in this cohort.

No association with change in HRQOL, measured via GIS UTN, was seen between either age, sex, being non-Caucasian, BMI, IMD, or married or cohabiting (Table 8.6). However, attendance at further education was associated with an improvement in HRQOL in this cohort. There was also an association between the interaction between the alcohol category 'daily or almost daily' with time and HRQOL. This association indicated that the relationship of alcohol consumption and HRQOL, measured via UTN, changed over time. However, no association between time and the other alcohol consumption categories were observed.

Time was not associated with an improvement in HRQOL measured by GIS UTN after adjustment for gout-specific, comorbid, socio-demographic factors, BMI and alcohol frequency (Table 8.6).

GIS Wellbeing during an attack

A history of oligo/polyarticular flares, and allopurinol use were associated with a deterioration in HRQOL, measured using the GIS Wellbeing during an attack (GIS WBDA) subscale, over three years in this cohort (Table 8.6). An improvement in GIS WBDA was associated with experiencing a current flare. The frequency of gout flares or disease duration were not associated with change in GIS WBDA.

Greater pain and worse depression were associated with a deterioration in HRQOL (Table 8.6). No associations between change in HRQOL, measured via GIS WBDA, and the number of comorbidities, eGFR <60, the presence of body pain, or anxiety severity were observed.

An improvement in HRQOL, measured via GIS WBDA, was associated with older age in this cohort (Table 8.6). Reporting daily or almost daily alcohol consumption, was also associated

with an improvement in HRQOL, however no associations were observed with the other alcohol categories. No associations with change in HRQOL and sex, being non-Caucasian, BMI, IMD, married or cohabiting, or further education were observed.

When adjusted for gout-specific, comorbid, socio-demographic factors, BMI and alcohol frequency time was associated with an improvement in HRQOL measured by GIS WBDA (Table 8.6).

GIS Concern during an attack

Over three years the frequency of gout flares, and a history of oligo/polyarticular flares were associated with a deterioration in HRQOL, measured via the GIS Concern during an attack (GIS CDA) subscale (Table 8.6). Longer disease duration was associated with an improvement in HRQOL. Change in GIS CDA was not associated with either experiencing a current flare or using allopurinol in this cohort.

A deterioration in HRQOL over three years, measured via GIS CDA, was associated with greater pain, worse depression, and worse anxiety (Table 8.6). There was also an association with the interaction between GAD-7 score at baseline with time and HRQOL. This association indicated that the relationship between anxiety and HRQOL may change over time. No associations between change in HRQOL and the number of comorbidities, eGFR <60, or the presence of body pain were observed.

Older age was associated with an improvement in HRQOL, measured via GIS CDA, over three years (Table 8.6). The interaction between being non-Caucasian and time was associated with a deterioration in HRQOL. This association indicated that the relationship between being non-Caucasian and HRQOL may change over time. Sex, BMI, IMD, being married or cohabiting, further education or alcohol consumption were not associated with change in HRQOL measured via GIS CDA.

After adjustment for gout-specific, comorbid, socio-demographic factors, BMI and alcohol frequency, time was not associated with an improvement in HRQOL measured by GIS CDA (Table 8.6).

8.4.3 Factors associated with change in generic HRQOL

SF-36 PF10

The following gout-specific factors were associated with a deterioration over three years in generic HRQOL, measured via the SF-36 PF10: experiencing two flares, four flares, or five or more flares, and allopurinol use (Table 8.7). No associations were observed between change in HRQOL and experiencing a current flare, a history of oligo/polyarticular flares, or disease duration (Table 8.7).

A higher total number of comorbidities, eGFR <60, greater pain, the presence of body pain, and worse depression were associated with a deterioration in HRQOL, measured via SF-36 PF10, over three years (Table 8.7). No association between anxiety and change in SF-36 PF10 was observed in this cohort.

Older age and being female were associated with a deterioration in HRQOL, measured via SF-36 PF10, over three years (Table 8.7). Living in a less socioeconomically deprived area (IMD) and consuming alcohol (daily or almost daily, 3-4 times/week, 1-2 times/week and special occasions) were associated with an improvement in generic HRQOL (Table 8.7). There was an association between the interaction between IMD and HRQOL (Table 8.7). This association indicated that the relationship between IMD and HRQOL changed over time. No associations between change in HRQOL and being non-Caucasian, BMI, married or cohabiting, or attendance at further education were observed.

When adjusted for gout-specific, comorbid, socio-demographic factors, BMI and alcohol frequency time was not associated with an improvement in generic HRQOL measured by the SF-36 PF10 (see Table 8.7).

Table 8.7 Factors associated with generic HRQOL; results from SF-36 PF10 LMM with random intercept, and results from HAQ-DI LMM with random intercept and random slope with unstructured covariance

	SF-36 PF10 n=554 (β (95% CI))	**HAQ-DI** n= 594 (β (95% CI))
Intercept	**123.97(112.16, 135.77)**	**-0.89(-1.25, -0.53)**
Time	0.10(-0.00, 0.20)	-0.00(-0.00, 0.00)
Gout-specific		
0 gout flares†◊	reference	reference
1 gout flare†◊	-0.33(-2.19, 1.53)	-0.00(-0.04, 0.03)
2 gout flares†◊	**-2.44(-4.58, -0.30)**	0.02(-0.03, 0.07)
3 gout flares†◊	-0.14(-2.45, 2.18)	0.03(-0.05, 0.11)
4 gout flares†◊	**-3.76(-6.83, -0.69)**	0.02(-0.07, 0.11)
≥5 gout flares†◊	**-6.07(-9.49, -2.65)**	0.02(-0.06, 0.10)
Current flare◊	1.40(-1.30, 4.10)	0.01(-0.06, 0.07)
Oligo/polyarticular flares◊	-1.46(-2.99, 0.08)	**0.04(0.01, 0.08)**
Allopurinol use◊	**-2.34(-4.33, -0.35)**	0.04(-0.01, 0.09)
Disease duration	0.07(-0.03, 0.17)	-0.00(-0.01, 0.00)
Comorbidities		
Number of Comorbidities	**-1.15(-2.11, -0.19)**	0.03(-0.01, 0.06)
eGFR <60mL/min/1.73m^2	**-3.39(-6.34, -0.46)**	**0.10(0.01, 0.19)**
Pain NRS score◊	**-1.44(-1.89, -1.00)**	**0.05(0.04, 0.06)**
Body pain	**-3.17(-5.73, -0.61)**	**0.09(0.02, 0.17)**
PHQ-9 score	**-1.56(-2.00, -1.11)**	**0.05(0.03, 0.06)**
GAD-7 score	-0.10(-0.57, 0.37)	0.01(-0.01, 0.02)
Socio-demographic		
Age	**-0.57(-0.69, -0.45)**	**0.02(0.01, 0.02)**
Female	**-8.74(-12.74, -4.74)**	**0.17(0.03, 0.30)**
Non-Caucasian	-5.45(-14.99, 4.09)	0.02(-0.33, 0.38)
IMD	**0.72(0.23, 1.22**)	**-0.02(-0.03, -0.00)**
IMD x Time	**-0.02(-0.04, -0.00)**	**0.00(0.00, 0.00)**
Married or cohabiting	0.46(-2.42,3.34)	0.00(-0.90, 0.10)
Further education	1.41(-1.10, 3.92)	-0.05(-0.12, 0.03)
BMI & alcohol		
BMI◊	-0.09(-0.31, 0.12)	**0.01(0.00, 0.01)**
Alcohol consumption- Never	reference	reference
Special occasions	**6.14(0.09, 12.18)**	-0.14(-0.34, 0.07)
1-3 times/month	5.37(-0.61, 11.34)	**-0.25(-0.44, -0.05)**
1-2 times/week	**6.68(1.30, 12.05)**	**-0.29(-0.47, -0.11)**
3-4 times/week	**8.45(3.05, 13.84)**	**-0.31(-0.49, -0.13)**
Daily or almost daily	**8.54(3.15, 13.93)**	**-0.30(-0.49, -0.12)**

β linear mixed model coefficient **†** in previous 12 months at baseline, 12 months and 36 months, in previous 6 months at 6 and 12 months **◊Time-varying covariate** in questionnaire at baseline, 6, 12, 24 and 36 months

Number of comorbidities total number of self-reported comorbidities in questionnaire; **eGFR** <60 mL/min/1.73m^2 indicative of chronic kidney disease; **NRS pain in last week** ranges from 0 (no pain) to 10 (pain as bad as it can be)**; Body pain** (including ache or discomfort or stiffness) for one day or longer in the 4 weeks prior to baseline; **PHQ-9 depression score** ranges from 0 (minimal) to 27 (severe); **GAD-7 anxiety score** ranges from 0 (minimal) to 21 (severe); **BMI** body mass index kg/m^2; **IMD** Indices of multiple deprivation highest score indicates least deprived.

SF-36 PF10 scored from **0 to 100**; higher score indicating performs all types of physical activities including the most vigorous without limitations due to health, **positive β** coefficient = improvement, and **negative β** coefficient = deterioration.

HAQ-DI scored from **0 to 3**; higher score indicating greater activity limitation, **positive β** coefficient = deterioration, and **negative β** coefficient = improvement.

HAQ-DI

A history of oligo/polyarticular gout flares was associated with worse activity limitation, measured via the HAQ-DI, over three years (Table 8.7). The following gout-specific factors were not associated with change in activity limitation: gout flare frequency, experiencing a current flare, allopurinol use, or a longer disease duration (Table 8.7).

Worse activity limitation, measured using the HAQ-DI, was associated with eGFR <60, greater pain, the presence of body pain, and worse depression (Table 8.7). No association between the number of comorbidities or anxiety, and change in activity limitation were observed (Table 8.7).

As shown in Table 8.7 older age, female sex, and a higher BMI were associated with worse activity limitation, measured via HAQ-DI, over three years. However, an improvement in activity limitation was associated with living in a less socioeconomically deprived area (based on IMD) and alcohol consumption (1-3/month, 1-2/week, 3-4/week and daily or almost daily) (Table 8.7). No associations between change in activity limitation and the following socio-demographic factors were observed: being non-Caucasian, married or cohabiting, and attendance at further education.

Time was not associated with a change in activity limitation measured by the HAQ-DI when adjusted for gout-specific, comorbid, socio-demographic factors, BMI and alcohol frequency (see Table 8.7).

8.4.4 Checking the assumptions of linear mixed models

Analysis of residuals and BLUPs

Figure 8.2 to Figure 8.15 display the Q-Q plots, histograms, scatter plots of GIS subscales, SF-36 PF10 and HAQ-DI LMM residuals and BLUPs. The Q-Q plots and histograms of the standardised residuals for each of the HRQOL do not display an obvious left or right skew. In

the scatterplots of the residuals for each HRQOL model plotted against fitted values the variance of the residuals is not increasing or displaying a 'fanning out' pattern, which would have suggested heteroskedasticity i.e. a lack of uniformity in the variance (Bland, 2015). The Q-Q plots and histograms of the BLUPs for each of the HRQOL outcome do not display an obvious left or right skew.

As the raw SF-36 PF10 and HAQ-DI scores demonstrated a left and right skew respectively, a sensitivity analysis was undertaken to transform the raw SF-36 PF10 and HAQ-DI scores. Log, cubic, squared, and square root transformations were undertaken for the SF-36 PF10 and HAQ-DI scores, however these transformations did not improve the distribution plots of the LMM residuals or BLUPs.

Results from the analysis of GIS CO LMM residuals and BLUPs

Residuals

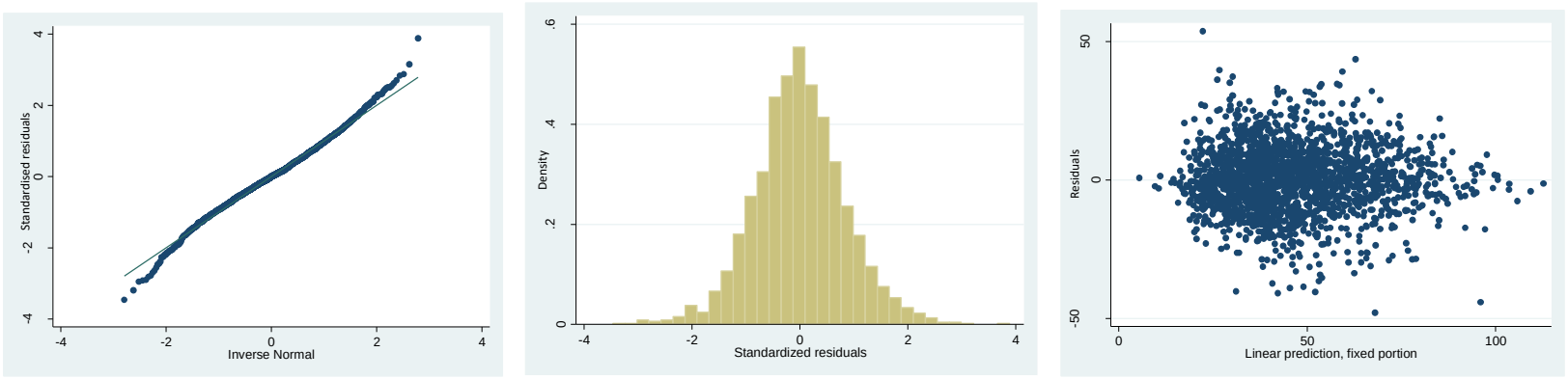


Figure 8.2 Q-Q plot, histogram and scatter plot of GIS CO residuals

Random Slope BLUPs

Random Intercept BLUPs

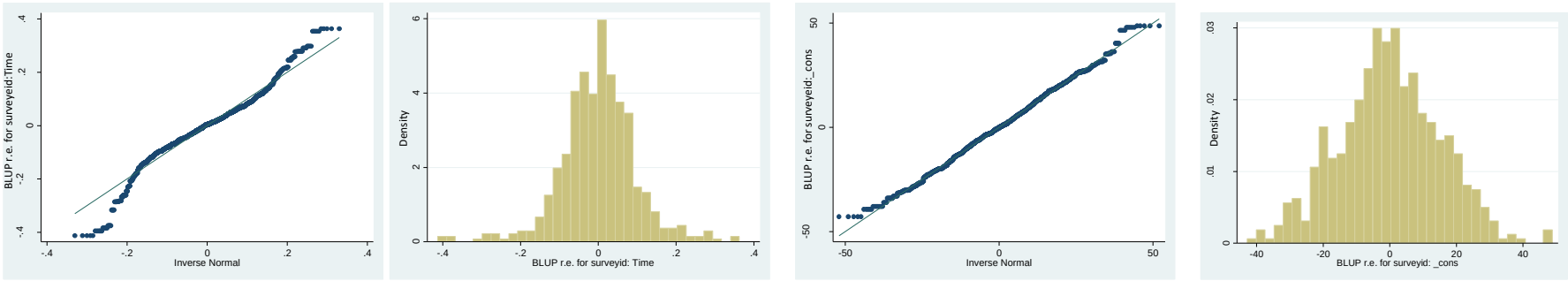


Figure 8.3 Q-Q plots and histograms of GIS CO BLUPs

Results from the analysis of GIS MSE LMM residuals and BLUPs

Residuals

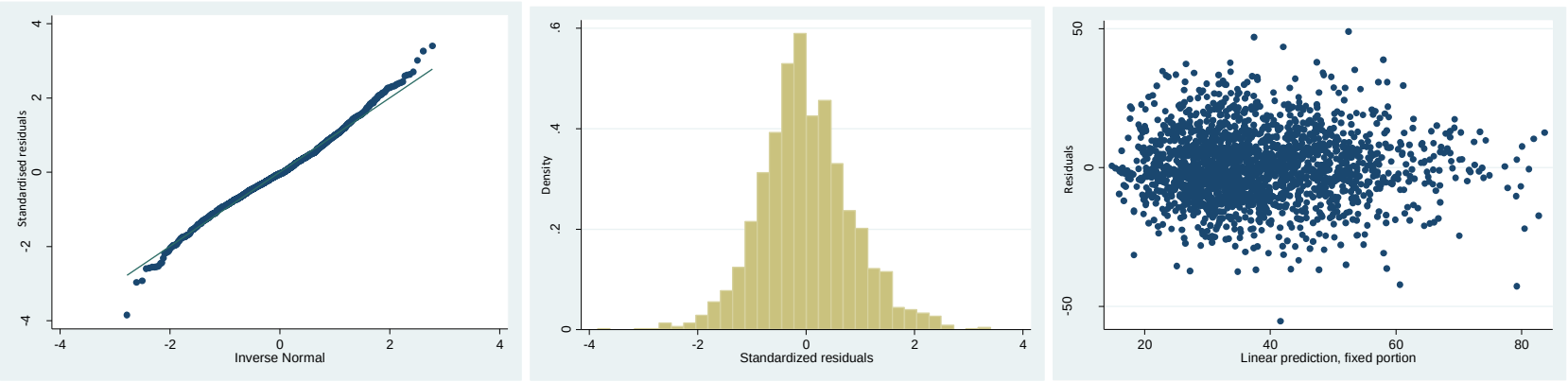


Figure 8.4 Q-Q plot, histogram, and scatter plot of GIS MSE residuals

Random slope BLUPs **Random intercept BLUPs**

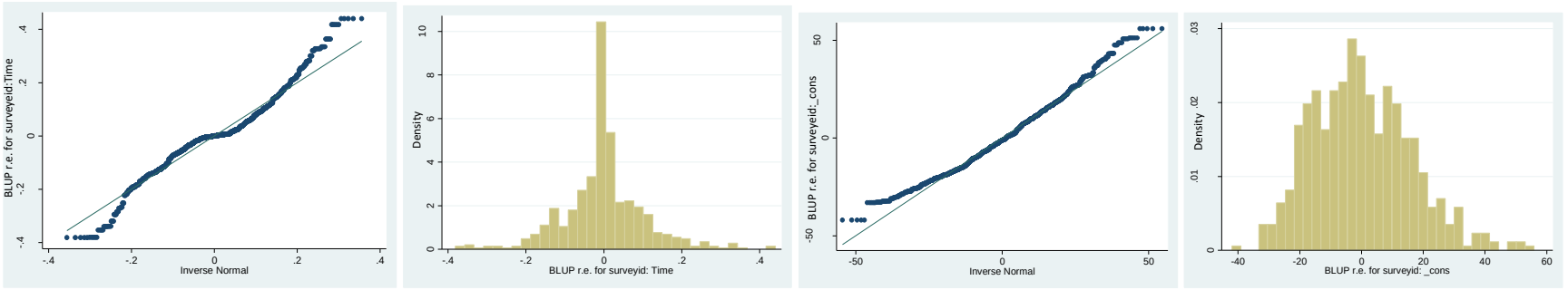


Figure 8.5 Q-Q plots and histograms of GIS MSE BLUPs

Results from the analysis of GIS UTN LMM residuals and BLUPs

Residuals

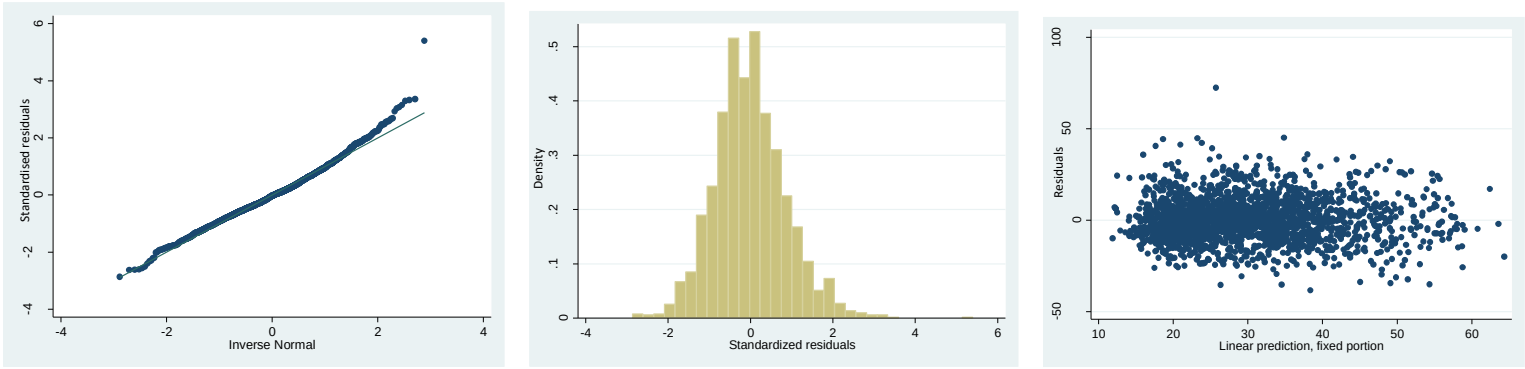


Figure 8.6 Q-Q plot, histogram, and scatter plot of GIS UTN residuals

Random Slope BLUPs

Random intercept BLUPs

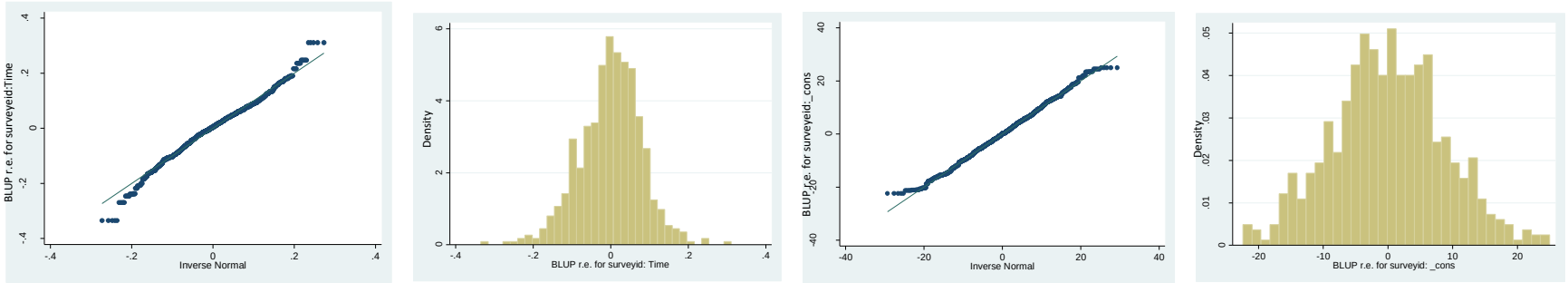


Figure 8.7 Q-Q plots and histograms of GIS UTN BLUPs

Results from the analysis of GIS WBDA LMM residuals and BLUPs

Residuals

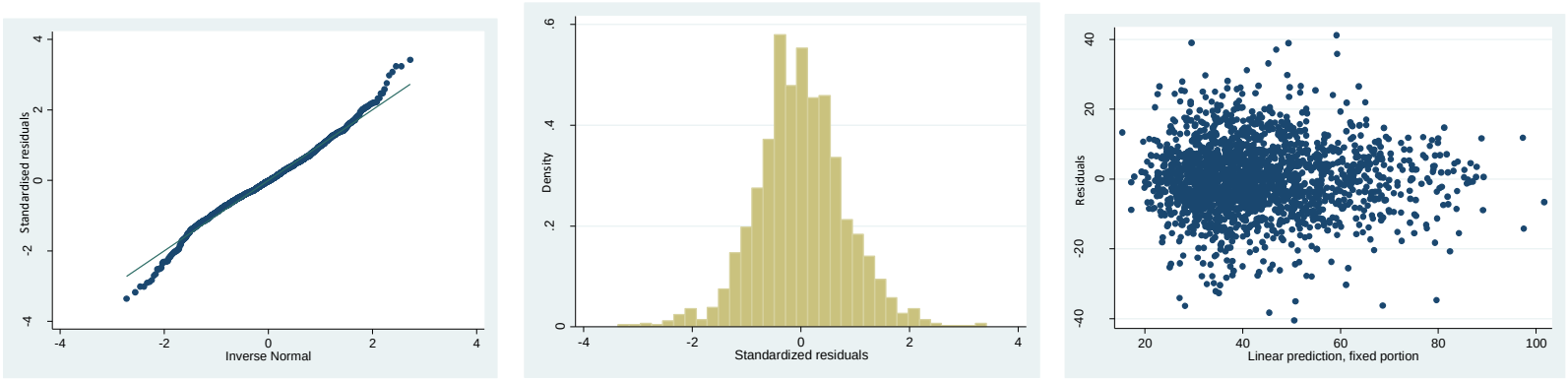


Figure 8.8 Q-Q plot, histogram, and scatter plot of GIS WBDA residuals

Random Slope BLUPs **Random intercept BLUPs**

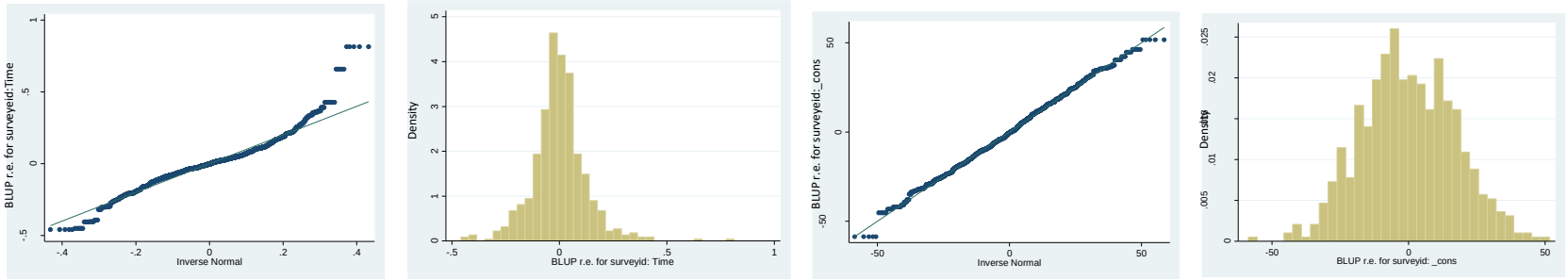


Figure 8.9 Q-Q plots and histograms of GIS WBDA BLUPs

Results from the analysis of GIS CDA LMM residuals and BLUPs

Residuals

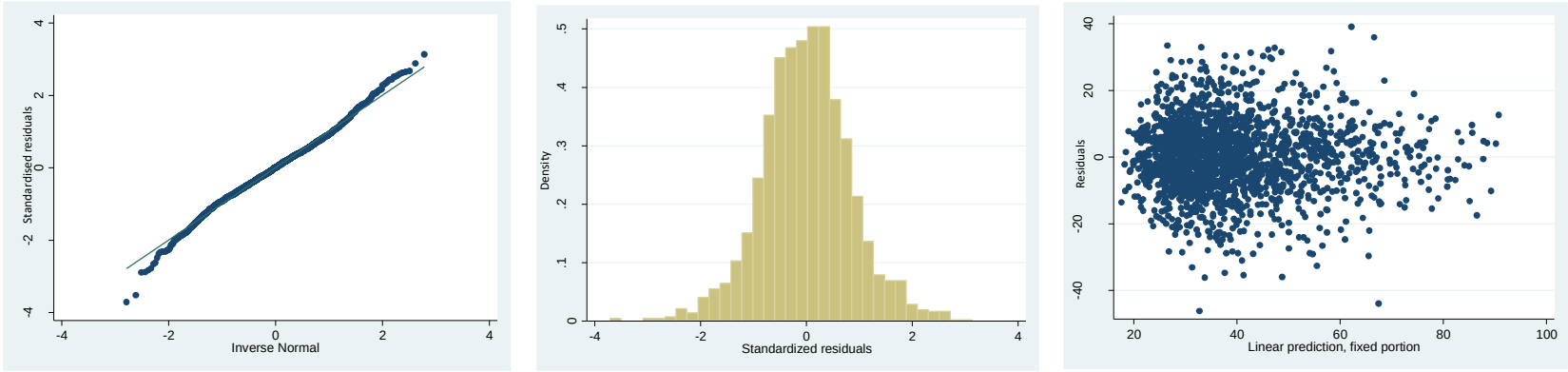


Figure 8.10 Q-Q plot, histogram, and scatter plot of GIS CDA residuals

Random slope BLUPs **Random intercept BLUPs**

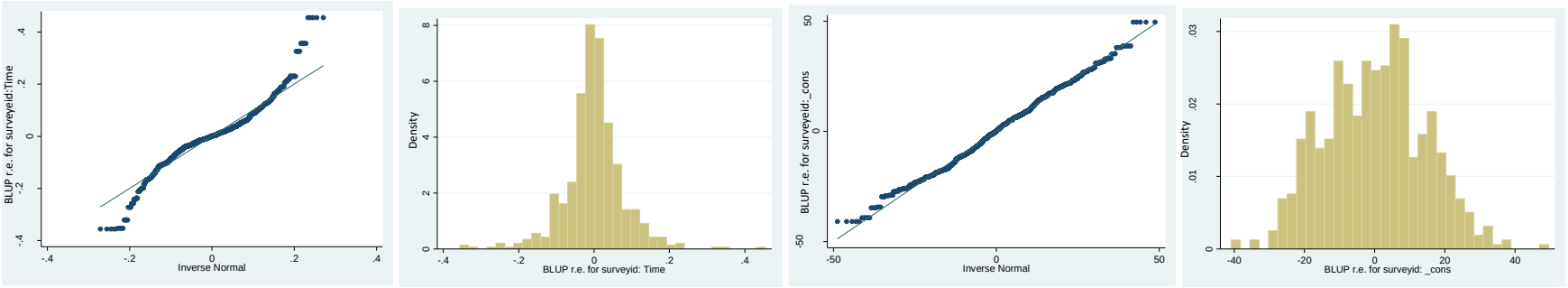


Figure 8.11 Q-Q plots and histograms of GIS CDA BLUPs

Results from the analysis of SF-36 PF10 LMM residuals and BLUPs

Residuals

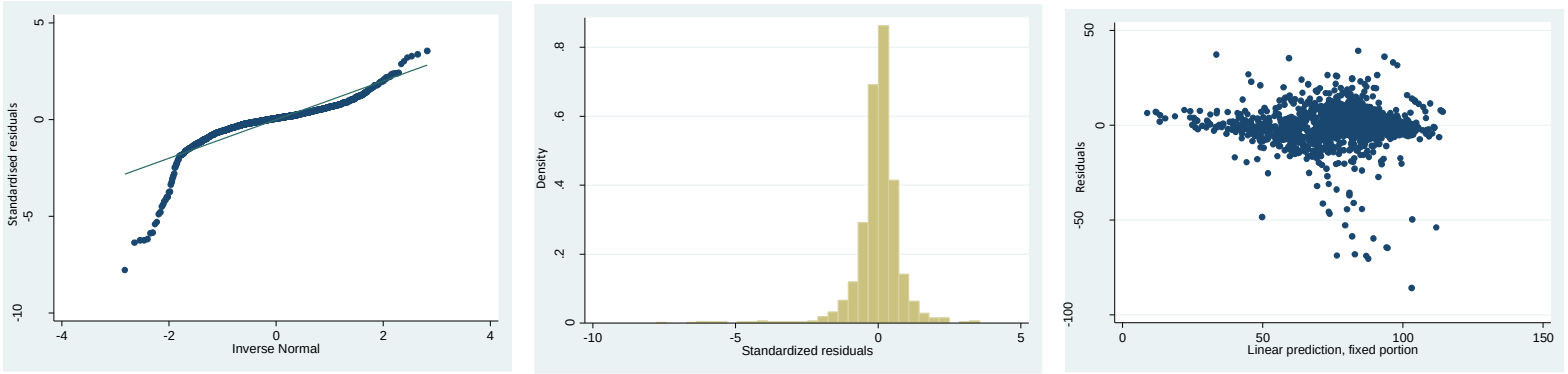


Figure 8.12 Q-Q plot, histogram, and scatter plot of SF-36 PF10 residuals

Random intercept BLUPs

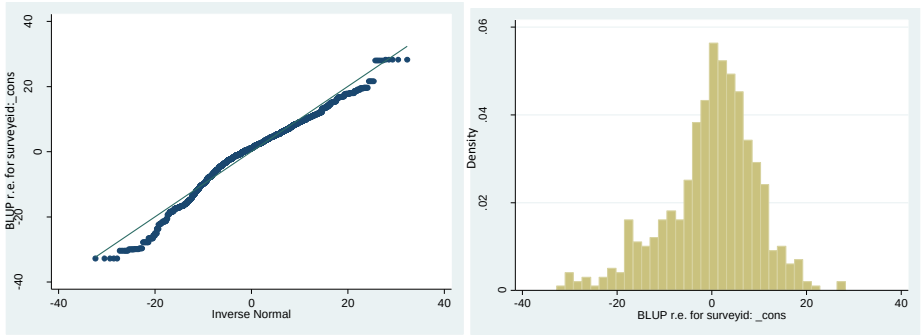


Figure 8.13 Q-Q plots and histograms of SF-36 PF10 BLUPs

Results from the analysis of HAQ-DI LMM residuals and BLUPs

Residuals

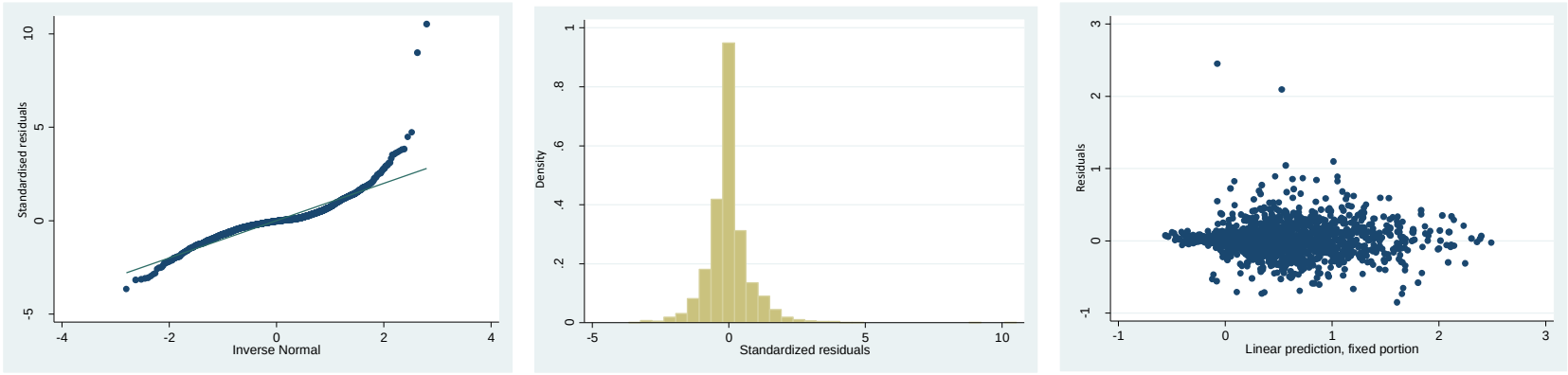


Figure 8.14 Q-Q plot, histogram, and scatter plot of HAQ-DI residuals

Random slope BLUPs **Random intercept BLUPs**

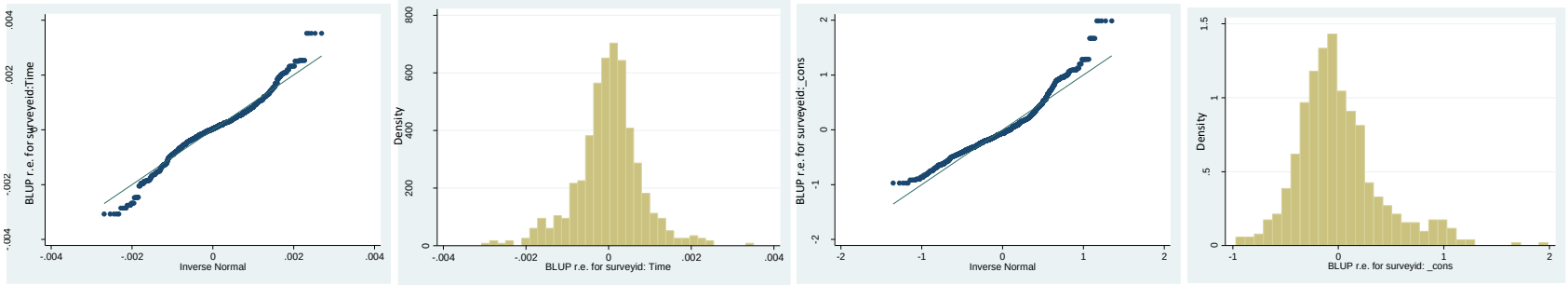


Figure 8.15 Q-Q plots and histograms of HAQ-DI BLUPs

8.5 Discussion

This chapter described the gout-specific, comorbid, socio-demographic and other factors associated with change in disease-specific HRQOL (measured via GIS subscales) and generic HRQOL (measured by SF-36 PF10 and HAQ-DI) over three years, in people living with gout in primary care.

8.5.1 Gout-specific factors associated with change in HRQOL

Gout flares were associated with a deterioration in disease-specific HRQOL over three years, with a dose-response relationship observed; the magnitude of deterioration in GIS CO, GIS MSE, and GIS UTN increased with the number of gout flares reported. This association between gout flare frequency and GIS subscale scores is consistent with the findings of previous cross-sectional studies (Chandratre et al 2018; Hirsch et al 2008) but was not investigated in previous prospective cohort studies. However, an association between change in total GIS score (the mean of all five GIS scales) over one year and physician rated gout severity score was identified by Wallace et al (2016) in a prospective US Veteran's medical centres cohort.

Gout flares were also associated with a deterioration in generic HRQOL, measured via the SF-36 PF10, which is in keeping with the association between the number of gout flares experienced and worse baseline SF-36 PF10 scores reported in cross-sectional studies previously (Becker et al 2009; Chandratre et al 2018). The lack of association between gout flares and activity limitation, measured via the HAQ-DI was also reported by Becker et al (2009).

Experiencing gout flares in more than one joint was associated with a deterioration in HRQOL in all of the GIS subscales (except GIS UTN) and in activity limitation (HAQ-DI). Previous cross-

sectional analyses have reported associations between worse disease-specific HRQOL, measured by the GIS subscales, and the number of joints involved in a gout flare (Hirsch et al 2010; Wallace et al 2016). Wallace et al (2016) found that baseline HAQ-DI scores were highly correlated with the number of joints affected by gout, and both Becker et al (2009) and Alvarez-Hernadez et al (2008) reported associations between baseline HAQ-DI scores and the number of tender, swollen, or painful joints.

Experiencing a current gout flare was associated with an improvement over three years in HRQOL measured by GIS WBDA. It might be expected that experiencing a current flare would be associated with worse HRQOL, rather than an improvement in HRQOL. A potential explanation for this finding relates to the time frame, 'during your last attack', used for questions in the GIS WBDA subscale. Some participants may be experiencing a current flare when completing the questionnaire but perhaps providing responses relating to their wellbeing during a previous flare. In contrast, experiencing a current flare was associated with a deterioration in the GIS CO and GIS UTN over the three years, indicating greater gout concern and unmet treatment need, which is consistent with the findings in previous cross-sectional studies (Chandratre et al 2018; Khanna et al 2015).

Allopurinol use was found to be associated with an improvement in HRQOL, measured by GIS UTN, indicating less unmet treatment need. Better HRQOL, measured by both GIS CO and UTN, over two years has also been reported in participants with greater ULT adherence in an RCT by Doherty et al (2018). The findings of this quantitative analysis are congruent with the experiences of people living with gout described in qualitative studies, as the perceived benefits of using allopurinol on HRQOL have been described in focus groups including participants from this cohort study (Chandratre et al 2016). Allopurinol was included as a time-varying covariate in the analysis in this chapter and was associated with a deterioration in HRQOL (GIS WBDA, SF-36 PF10) and activity limitation (HAQ-DI). However, a lack of

association with both generic HRQOL or disability, and allopurinol use has been reported by several cross-sectional analyses (Alvarez-Negegyei et al 2005; Chandratre et al 2018; Roddy, Zhang & Doherty, 2007c; Scire et al 2013). The association between allopurinol use and worse HRQOL observed in this cohort may be because participants with worse physical function, more severe comorbidities or gout symptoms impacting on wellbeing, were more likely to either be prescribed or take allopurinol. Consequently, this may indicate confounding by severity or indication (Szklo & Javier, 2014). Participants in both focus groups and phone interviews have cited frequent flares, or the fear of recurrent flares or unpredictable flares, as reasons for starting allopurinol (Chandratre et al 2016; Harrold et al 2010). These comments from qualitative studies fit with the notion that people with worse gout may be more likely to take allopurinol. In addition, as allopurinol can induce flares through a reduction in urate levels, due to initiation of allopurinol or an increase in allopurinol dose (Seth et al 2014), it is perhaps understandable that participants who reported to taking allopurinol would experience a deterioration in feelings of wellbeing during an attack over time.

Longer disease duration at baseline was associated with an improvement in disease-specific HRQOL in all GIS subscales with the exception of WBDA GIS subscale. This could in part be attributable to the process of adaption. Adapting to an illness over time and learning to cope are acknowledged aspects of living with a chronic illness which can influence how a person reflects on their HRQOL (Fayers & Machin, 2016). The lack of association between disease duration and GIS WBDA, SF-36 PF10 or HAQ-DI suggests that living with gout for a longer period of time does not improve participants' wellbeing during a flare or their generic HRQOL over time.

8.5.2 Comorbidities associated with change in HRQOL

A higher number of self-reported comorbidities at baseline was associated with a

deterioration in generic HRQOL, measured via SF-36 PF10. This result is consistent with the cross-sectional findings of Becker et al (2009) who reported worse SF-36 PF10 scores in participants with one or more comorbidities, compared to no comorbidities, and Scire et al (2013) who reported an association between the number of self-reported comorbidities and worse SF-36 PCS scores. However, no association between disease-specific HRQOL, measured via GIS, was identified. Wallace et al (2016) reported a lack of association between the Charlson comorbidity index and change in total GIS over one year.

Poorer generic HRQOL compared to controls has been reported in a general cohort of individuals with chronic kidney disease (CKD) at all stages of CKD (Pagels et al 2012), this is consistent with having an eGFR less than 60 mL/min/1.73m^2 being associated with a deterioration in generic HRQOL and a worse activity limitation in people with gout in this cohort. Worse generic HRQOL and activity limitation was also associated with self-reported renal failure in this cohort at baseline (Chandratre et al 2018) and a correlation between the PCS of the SF-36 and renal failure was reported in a cross-sectional analysis in people with gout by Lee et al (2009).

Both the severity and presence of pain were associated with a deterioration in disease-specific and generic HRQOL over three years in this cohort. Worse pain VAS scores have been shown to associate with the GIS subscale and SF-36 PF10 scores cross-sectionally (Becker et al 2009; Hirsch et al 2010) and both severity and presence of pain with activity limitation in prospective studies (Alvarez-Hernadez et al 2008; Stewart et al 2018). The association between worse HRQOL and pain is in keeping with qualitative studies as pain intensity has been highlighted in semi-structured interviews as something which people with gout find particularly worrying (Liddle et al 2015).

A greater depression severity score (PHQ-9) at baseline was associated with a deterioration in

disease-specific HRQOL (GIS MSE, WBDA and CDA), generic HRQOL (SF-36 PF10) and activity limitation (HAQ-DI). Cross-sectional associations between depression and worse GIS, SF-36 PF10 and HAQ-DI scores have been reported previously (Chandratre et al 2018). In a review of systematic reviews, Sivertsen et al (2015) also described associations between severity of depression and poorer generic HRQOL in older persons (without gout) in both cross-sectional and longitudinal studies. It has been postulated that this may partly be due to poorer coping strategies in individuals with depressive disorders (Sivertsen et al 2015), whilst other authors have highlighted that depression is associated with an increased risk of non-adherence to treatment in people living with chronic diseases (Grenard et al 2011).

A more severe generalised anxiety score (GAD-7) at baseline was associated with a deterioration in disease-specific HRQOL but not generic HRQOL in this cohort. This finding is understandable as the association with anxiety scores was with subscales which relate to 'concern' (GIS CO or GIS CDA) or have items which explicitly refer to 'worry' (GIS MSE). Conversely no association between GAD-7 and either SF-36 PF10 or HAQ-DI scores over three years was identified in this cohort. Alvarez-Hernadez et al (2009) also reported no correlation between anxiety and HAQ-DI scores in their cohort study. However, an association between anxiety and lower SF-36 PF10 scores have been reported in this cohort at baseline (Chandratre et al 2018) and also in a general primary care cohort in the US (Brenes, 2007). However, these were cross-sectional analyses of the relationship between anxiety and generic HRQOL, whereas this analysis investigated change over time.

8.5.3 Socio-demographic factors associated with change in HRQOL

Older age at baseline was associated with an improvement in disease-specific HRQOL in all of the GIS subscales, except UTN, which is in keeping with Wallace et al (2016) who described an improvement in total GIS scores over one year for older participants. The association between

age and better disease-specific HRQOL could be attributable to several factors. Older participants are less likely to be working therefore may have responded more favourably to items relating to the impact of gout on work. Also, as social interaction can often decline as adults age (Cudjoe et al 2020), the impact of gout on social activities may be perceived as less important for older members of the cohort. Older participants may have also responded with a different internal rating scale, due to differing levels of stoicism, in comparison to younger participants (Moore et al 2013; Sivertsen et al 2015).

Conversely, older age at baseline was associated with both a deterioration in generic HRQOL and worse activity limitation. Several cross-sectional studies have reported an association between age and worse HRQOL in the physical domain (Chandratre et al 2018; Lee et al 2009; Roddy, Zhang & Doherty, 2007c). Deterioration in generic HRQOL in this older cohort could have perhaps been expected considering that the likelihood of experiencing complex health conditions, or being disabled, increases with age (Office for National Statistics, 2018). A decline in HRQOL with increasing age has been identified in another longitudinal analysis of an older primary care cohort (Eisele et al 2015).

Being female was associated with both a deterioration in generic HRQOL and worse activity limitation. Although these findings are consistent with the association identified at baseline between being female and both worse SF-36 PF10 and HAQ-DI scores (Chandratre et al 2018), other authors have reported no association between sex and either activity limitation (Alvarez-Hernadez et al 2008) or HRQOL in the physical domain (Roddy, Zhang & Doherty, 2007c). However, the association observed in this prospective cohort study is in keeping with, a review of longitudinal (non-gout) population data for the SF-36 which showed that females experience greater declines in HRQOL in comparison to men (Obidoa, Reisine & Cherniack, 2010).

Living in a less socioeconomically deprived area, indicated by IMD score, was associated with an improvement in generic HRQOL and less activity limitation. These findings are consistent with the association identified at baseline between worse deprivation and both worse SF-36 PF10 and HAQ-DI scores (Chandratre et al 2018) and the association between measures of deprivation and change in physical health identified in longitudinal cohorts of older people (without gout) (Myck, Najsztub & Oczkowska, 2019). Attendance at further education was associated with an improvement in disease-specific HRQOL GIS UTN. This improvement in HRQOL on the unmet treatment domain may be attributable to a range of factors. Individuals with higher levels of educational attainment may have developed better problems solving skills and be more able to adapt to the challenges of a change in health status (Edgerton, Roberts & von Below, 2012). Also, higher educational attainment has been associated with greater concordance to medication (Mathes, Jaschinski & Pieper, 2014).

8.5.4 Strengths and limitations

For the first time the factors associated with change in disease-specific and generic HRQOL in people living with gout in primary care over three years have been described.

Previous research, clinical knowledge, and findings from earlier chapters in this thesis were used to select the covariates included in the global model. This use of background knowledge to generate covariates for inclusion in the model has been advocated as good practice (Cheng et al 2010; Harrison et al 2018; Heinze, Wallsich & Dunkler, 2018; Long, 2012). It is important to acknowledge that not all covariates included in this analysis were time-varying covariates. Thus, interactions between time-invariant covariates and time were included in the LMMs initially to test whether the effect of these covariates on change in HRQOL altered with time (Berrington, Smith & Sturgis, 2006; Long, 2012). The use of backward selection has been advocated in preference to forward deletion in the selection of covariates in a model (Cheng

et al 2010; Heinze, Wallsich & Dunkler, 2018), thus this method was used to identify interactions in the LMMs which could be removed. Several interactions were retained in the model for this exploratory analysis, as they suggested that the effects of these covariates on specific HRQOL outcomes changed over time. Previous authors investigating the factors associated with change in HRQOL in people with gout have reduced models further to remove covariates which did not yield a statistically significant association with GIS scores (Wallace et al 2016). However, covariates should not be removed from a model if theory suggests they may be relevant to the outcome, or they may be a potential confounder in the relationship between another covariate and the outcome (Cheng et al 2010; Heinze, Wallsich & Dunkler, 2018). Consequently, further backward deletion to remove covariates was not undertaken.

The potential for unmeasured confounding in this model should be acknowledged, however the optimal method to account for unmeasured confounding specifically in longitudinal data and time-varying covariates has yet to be determined (Streeter et al 2017). The inclusion of a range of covariates relevant to disease-specific and generic HRQOL in the LMMs aimed to reduce the potential for unmeasured confounding in these models**.** However, tophi and serum urate were not included as covariates in the LMMs due to the low prevalence of tophi (2.1%) and the high proportion of missing serum urate levels (missing 61.1%). The presence of tophi has been associated with worse HRQOL measured via HAQ-DI and SF-36 in cross-sectional analyses in people living with gout (Alvarez-Negegyei et al 2005). Minimal correlations have been reported between the presence of tophi and GIS CO and GIS UTN and between serum urate and GIS CO in a cross-sectional study by Hirsch et al (2010). However, no correlation between serum urate level and the GIS subscales was reported in a prospective study by Wallace et al (2016).

To improve the accuracy of the results of this analysis, the statistical methods deployed

specifically dealt with the serial correlation, variance, and covariance in the longitudinal HRQOL outcomes (Long, 2012; West, Welsh & Galecki, 2015). As the variance component and standard errors of multilevel models can be biased when the outcome is skewed (Maas & Hox, 2004), the SF-36 PF10 and HAQ-DI were modelled in this chapter using robust standard errors. In simulation studies involving data with a skewed distribution, this use of robust standard errors with multilevel models has been shown to lead to less bias in standard errors compared with when standard multilevel models are used (Zhu & Gongalez, 2017). The analysis of the residuals for each of the LMMs in this chapter, to investigate the assumptions of normality and constant variance, revealed acceptable results.

As discussed in chapter four the outcomes and covariates included in the analyses in this chapter included a degree of missing data. Missing data in outcome measures are permitted in LMM (Cnaan, Laird & Slasor, 1997; Long, 2012; West, Welsh & Galecki, 2015). Imputation of missing data prior to linear mixed modelling has been deemed unnecessary and has been shown to not improve the precision of estimates (Peters et al 2012; Twisk et al 2013).

As described in the cohort study's protocol paper (Chandratre et al 2012) the *a priori* sample size for this analysis was 882. This was based on detecting a smallest meaningful difference in HRQOL of 0.2 standard deviation units detected between two groups (less than 2 gout flares per year and two or more gout flares per year) at a significance of 0.05, power of 90% and autocorrelation of 0.8. This sample size calculation was based on an expected response rate of 70% at baseline and drop out over the three-year follow up of 30%. However, as described in chapter four the response rate at baseline was below this expected level and the proportion of participants dropping out over three-years was higher than the 30% anticipated. The samples sizes used in this thesis are discussed in the following discussion chapter.

8.5.5 Implications for clinical practice

Measurement of HRQOL has the potential to facilitate decision making concerning clinical management, service development and health policy (Green, Brazier & Deverill, 2000; Hahn et al 2007; Hennessy et al 1994; Fayers & Machin, 2016) through an understanding of the wider impact of disease on patients and identifying patient characteristics associated with poorer outcomes (Calvert & Freemantle, 2003; Fitzpatrick et al 1992). The identification of the factors associated with change in HRQOL in people living with gout in primary care is of particular relevance to clinical practice as the majority of patients are managed within primary care.

This research demonstrates the impact of experiencing gout flares on disease-specific HRQOL, with the deterioration in HRQOL increasing with the number of gout flares experienced. When one considers that even one flare was associated with a deterioration in disease-specific HRQOL in this cohort, this provides additional justification for the recommendation that ULT, should be discussed with all patients living with gout (Hui et al 2017). The deterioration of both disease-specific and generic HRQOL over time in participants with worse pain, the presence of body pain, and more severe depression could have implications for targeting interventions in clinical practice. In addition, the improvement over time in disease-specific HRQOL in participants who attended further education and the improvement in generic HRQOL in participants living in less deprived areas, highlights the disparity in change in HRQOL for those with different levels of educational attainment and deprivation. These findings could have implications for patient information and education, along with prioritisation of resources.

The implications of the findings in this chapter for clinical practice and people living with gout are discussed further in section 9.4.

8.6 Conclusion

In conclusion, this chapter describes the factors associated with change in gout-specific HRQOL and generic HRQOL scores over a three-year period in a prospective cohort study of people living with gout in primary care using linear mixed models. Prospective analysis of HRQOL, with the inclusion of time-varying covariates, enabled the change in HRQOOL in people living with gout to be captured. Factors associated with deterioration in both disease-specific and generic HRQOL included more frequent gout flares, a history of oligo/polyarticular flares, allopurinol use, the presence of body pain, worse pain severity, and worse depression score. Factors associated with improvement in disease-specific HRQOL included allopurinol use, a longer disease duration, older age, and attendance at further education, whilst factors associated with improvement in generic HRQOL included living in a less socioeconomically deprived area and consuming alcohol more frequently.

In the next chapter key findings of the whole thesis are summarised, strengths and limitations are discussed, and the implications of the whole thesis for both clinical practice and future research are considered.

9 Chapter Nine Discussion and conclusion

9.1 Overview of chapter and aim

In the previous chapter, the factors associated with change in HRQOL over a three-year period in a prospective cohort study of people living with gout in primary care were described.

The aims of the following chapter are to discuss the key findings of the whole thesis, considering the strengths, limitations, implications for clinical practice, implications for future research, and to draw conclusions.

9.2 Key findings of the thesis

The key findings in this chapter are derived from the analysis of outcomes from a three-year prospective cohort study in people living with gout in primary care.

Frequency and pattern of gout flares over three years

Chapter five described self-reported gout flares over time. At each time-point, the largest proportion of participants reported experiencing no gout flares, however over a quarter of the cohort reported two or more gout flares indicating frequent flares. Diverse individual patterns of change in gout flares over time were identified. Various approaches to modelling longitudinal gout flare data for the whole cohort were trialled using latent growth curve modelling (e.g. trialling ordinal, zero-inflated Poisson and negative binomial models with different functional forms) and more favourable results were returned for a quadratic ordinal latent growth curve model.

Gout flare trajectory classes

In chapter six, latent class growth analysis was used to identify and describe distinct gout flare trajectory classes over a three-year period. Six distinct trajectory classes were identified which displayed a range of patterns in gout flares over time: 'frequent and persistent', 'gradually

worsening', 'frequent then improving', 'moderately frequent', 'moderately frequent then improving' and 'infrequent'. The selection of a six-class solution was justified by the statistical fit and the clinical relevance of the trajectories. Frequent flare classes had more participants who were socioeconomically deprived, obese, had chronic kidney disease, and a history of oligo/polyarticular flares. Less frequent flare classes were associated with allopurinol use, and a lower serum urate level.

Change in HRQOL over three years and factors associated with change

Chapters seven and eight investigated the change in gout-specific HRQOL (GIS subscales) and generic HRQOL (SF-36 PF10 and HAQ-DI) over time and factors associated with change. At a group level, the magnitude of change in HRQOL scores was small and for some HRQOL scores negligible, however wide variation in change in HRQOL scores at an individual level was identified. HRQOL measures (GIS subscales, SF-36 PF10 and HAQ-DI) displayed a linear form and were correlated over time. The SF-36 PF10 and HAQ-DI demonstrated a left and right skew respectively.

Factors associated with deterioration in HRQOL over three years were gout flares (GIS CO, MSE, UTN, CDA, SF-36 PF10), current flare (GIS CO, UTN), history of oligo/polyarticular flares (GIS CO, MSE, WBDA, CDA, HAQ-DI), using allopurinol (GIS WBDA, SF-36 PF10), number of comorbidities (SF-36 PF10), eGFR <60mL/min/1.73m^2 (SF-36 PF10, HAQ-DI), worse pain severity (GIS CO, UTN, WBDA, CDA, SF-36 PF10 and HAQ-DI), the presence of body pain (GIS CO, MSE, SF-36 PF10 and HAQ-DI), worse depression score (GIS MSE, WBDA, CDA, SF-36 PF10 and HAQ-DI), worse anxiety score (GIS CO, MSE, CDA), older age (SF-36 PF10, HAQ-DI), being female (SF-36 PF10, HAQ-DI), and higher BMI (HAQ-DI). Factors associated with an improvement in HRQOL over three years were experiencing a current flare (GIS WBDA), using allopurinol (GIS UTN), longer disease duration (GIS CO, MSE, UTN, CDA), older age (GIS CO, MSE, WBDA, CDA), living in a less deprived area (SF-36 PF10, HAQ-DI), attendance at further

education (GIS UTN), and consuming alcohol more frequently (GIS WBDA, SF-36 PF10, HAQ-DI).

9.3 Strengths and Limitations

New knowledge

The analysis presented in this thesis is unique as it was undertaken in a large prospective cohort of over 1000 people living with gout in primary care and followed up for three years. This analysis of longitudinal outcome measures, as opposed to solely cross-sectional analysis, enabled the dynamic nature of these outcomes to be captured (Fayers & Machin, 2016; Hahn et al 2007).

This thesis uniquely identified groups of people reporting similar patterns of change in gout flares over time. Although trajectories of pain and disease activity have been studied previously in other rheumatological conditions, this is the first time that trajectories of gout flares and associated participant characteristics have been investigated.

There was a paucity of previous research into longitudinal change in both disease-specific and generic HRQOL in people living with gout specifically in primary care. This is the longest prospective study to investigate the factors associated with change in both disease-specific and generic HRQOL in people living with gout in primary care, whilst adjusting for a range of gout-specific, comorbid, socio-demographic and other factors.

Combining statistical analysis and clinical interpretation

The results of statistical analysis were not considered in isolation and the clinical interpretation of findings were considered throughout this thesis. For example, both the clinical interpretation of the pattern of gout flares over time and the characteristics of participants in the gout flares trajectory classes supported the selection of a model with six

gout flare trajectory classes. The findings have also been presented to and discussed with Rheumatologists, GPs and lay members of the PPIE Keele University Research User Group (RUG). Clinicians commented that the flare trajectories reflected patients encountered in clinical practice (experiencing worsening, improving and no change in symptoms), and the factors associated with change in HRQOL were plausible. PPIE group members were able to identify with the trajectory classes and were able to say which trajectory class they feel they would be assigned to based on their symptoms. PPIE group members were also able to relate to the factors associated with change in HRQOL identified in this study.

Planning statistical analysis

Decisions regarding the statistical analysis undertaken and criteria used to interpret results were informed by the literature in the field and specified *a priori* for each analysis. Time was dedicated to understanding the behaviour of the outcome measures over time, prior to undertaking the main analyses, due to the paucity of previous studies undertaking longitudinal modelling on these outcomes. This investigation then informed how these outcome measures should be statistically modelled. Consideration of the behaviour of outcomes prior to analysis is frequently omitted and can have detrimental consequences for interpretation (Feldman, Masyn & Conger, 2009). This led to the use of quadratic ordinal models for the LCGA of gout flares and the addition of robust standard errors to LMMs for the skewed HRQOL outcomes.

Attrition bias

The investigation in chapter four revealed that responders were more likely to be male and appeared to be younger, better educated and less deprived compared with non-participants. Responders also had less severe (or better treated) gout, were less unwell and had better quality of life. Thus, these findings provided evidence for attrition bias within this cohort. It was important to investigate attrition bias to better understand this cohort as this

thesis was the first time that prospective data from this cohort had been analysed. As participants in this study seemed to have less severe gout and better HRQOL it is important to consider that the findings of this thesis, such as the strength of the associations observed, may have been different if no attrition bias had been identified.

Missing data

The proportion of independent and dependent variable missing data on the questionnaires was below 10% for most variables, however missing data was identified in some dependent variables and covariates. The statistical modelling techniques used to analyse these outcomes, LCGA and LMM, were selected due to their ability to deal with missing data in outcome measures (Berlin, Parra & Williams, 2014; Cheng et al 2010; Jung & Wickrama, 2008; Locascio & Atri, 2011). In LMM multiple imputation of outcome measures has been shown to not improve the precision of estimates (Peters et al 2012), thus imputation was not undertaken prior to using LMMs. In the LCGA in this thesis the sensitivity analysis undertaken showed that a six-class gout flare trajectory model was selected (based on statistical indices and similar trajectories) regardless of the degree of missing data.

Case ascertainment

Participants were identified by a Read code for gout or a prescription for allopurinol or colchicine within the medical records in the two years prior to baseline. Thus, the diagnosis was made in primary care where the gold standard for gout diagnosis, the detection of MSU crystals in joints or tissues, is not routinely undertaken (Richette et al 2020). Whilst there is potential for misclassification bias, Read codes for gout have displayed a reported sensitivity of 93.7% and a positive predictive value of 97.6% (Hassey, Gerret & Wilson, 2001).

Outcome measures

Although a definition of gout flares has recently been validated for research participants with gout (Gaffo et al 2018), at the initiation of the cohort study there was a lack of a widely accepted valid definition for flares (Gaffo et al 2012). The validated flare definition by Gaffo et al (2018) requires information about clinical features (pain severity, warmth, swelling) of each flare but these clinical parameters were not collected within the questionnaire. Whilst flares self-reported in this cohort study may be liable to misclassification bias, they have been shown to have a sensitivity of 91% and negative predictive value of 96% when compared to the criterion of an assessment by a rheumatologist (Gaffo et al 2012). Alternative methods used to identify gout flares such as consultation or prescription data are likely to underestimate gout flare frequency, as flares are often self-managed by patients without seeking medical attention (Chandratre et al 2016; Macfarlane et al 2016; Neogi et al 2006; Rothenbacher et al 2011; Sarawate et al 2006).

Whilst the SF-36 has been endorsed by OMERACT as a measure of HRQOL in people living with gout (Singh et al 2011a), only the SF-36 PF10 was included in the questionnaires and available for analysis in this study. The inclusion of the SF-36 PF10 only was justified by existing evidence at that time that impairment of HRQOL was predominantly on the physical domain for people living with gout (Chandratre et al 2013) and reduced questionnaire length and therefore the burden to participants. The disease-specific HRQOL measures analysed, the GIS subscales, have yet to be endorsed by OMERACT as an outcome measure for people living with gout. When the GIS was considered by OMERACT in 2011 (Singh et al 2011a), it was not endorsed due to the panel's lack of familiarity with the scale and a lack of information regarding the feasibility of its use. The GIS remains the only disease-specific HRQOL measure for people living with gout (Janssens et al 2019) and it has been used to measure HRQOL in subsequent prospective cohort studies (Singh et al 2016; van Leeuwen 2018; Wallace et al

2016) and a RCT (Doherty et al 2018). This study uniquely undertook longitudinal analysis of all five subscales of the GIS, which address different aspects of disease-specific HRQOL. This use of the GIS prospectively adds to the limited knowledge regarding the behaviour of the GIS subscales over time. The responsiveness of the GIS scale, in relation to the effect sizes and standardised response means, have been reported previously over shorter periods of time, 8 weeks in a RCT (Khanna et al 2011b) and a prospective cohort over 12 months in secondary care (Wallace et al 2016). The report of effect sizes and SRMs over three years in a primary care cohort is therefore unique, thus provided additional insight into the responsiveness of this outcome measure.

Recall bias and measurement error

It is important to acknowledge the potential for recall bias, where participants have selective memories when recalling past events (Bowling, 2009), in relation to the reporting of both outcome measures and covariates in this study. Gout flares in this study were self-reported and hence at risk of recall bias however, flares are salient and memorable owing to their dramatic nature (Harrold et al 2010; Roddy, Mallen & Doherty, 2013; Taylor et al 2015; Teng, Nair & Saag, 2006). Using more frequent questionnaires or symptom diaries may have reduced recall bias but would have increased participant burden (Bowling, 2009; McColl, 2004).

Unmeasured covariates

Several gout-specific, comorbid, socio-demographic and other characteristics of participants were considered when investigating change in gout flares and HRQOL. However, some factors were not available for inclusion, for example dietary factors.

The inclusion of time-varying covariates enabled the dynamic relationship between covariates and HRQOL to be investigated. However, some covariates were only included in the HRQOL

models at baseline as they were not included in the questionnaire at every time point, for example comorbidities, body pain, depression and anxiety.

Response shift

A change in someone's assessment of their HRQOL when in fact there has been no objective change in their circumstances is an example of 'response shift', which may be due to several factors including changes in the internal scale participants use to rate themselves when responding to items (Blome & Augustin, 2015; Fayers & Machin, 2016) and change in participants' perceptions of their circumstances due to adaption or adjustment over time (Bowling, 2009; Blome & Augustin, 2015; Fayers & Machin, 2016). Response shift may be considered a form of measurement bias (Blome & Augustin, 2015; Fayers & Machin, 2016) and has been identified in prospective studies investigating HRQOL and in other clinical conditions (Barclay-Goddard et al 2011; Zhang et al 2012c). Whilst methods have been proposed to test for response shift e.g. 'thentest', they can be influenced by recall bias (Bloome & Augustine, 2015). Whether response shift is a form of bias or simply reflective of a person's ability to cope and adjust to their circumstances has been debated (Blome & Augustin, 2015).

Regression to the mean

It is prudent to consider the potential for regression to the mean; where extreme measurements at the upper or lower end of the measurement scale are likely to be followed by less extreme results (Bowling, 2009; Barnett, van der Pols & Dobson, 2005). The fact that this is a cohort study of prevalent cases rather than an inception or intervention study means that it is less likely participants were recruited at a time when their symptoms were more severe. However, prevalent cases may have been more likely to take part if their symptoms were worse at the time of invitation. The results from the LMMs without covariates do not

suggest that participants with poor HRQOL scores at baseline experienced better scores at later time-points. In fact, the negative covariance between intercept and slopes for some of the HRQOL scores (eg. GIS CO, GIS MSE, GIS UTN, and WBDA) indicated that the higher the GIS subscale score at baseline (intercept) the smaller the rate of change (Twisk, 2006).

Hawthorne effect

It is important to acknowledge the potential for a 'Hawthorne effect' amongst participants in this observational cohort, where the fact that participants are being observed could influence the outcomes reported (Bowling, 2009). The fact that participants were followed up over time may have partly led to the short-term reduction in gout flares reported by some participants, perhaps because participation in the study influenced their behaviour by increasing their awareness of the topic or simply because they felt someone was taking an interest in them (Bowling, 2009).

Sample size and power calculations

A sample size calculation was not undertaken prior to investigating gout flare trajectories as there are no set rules for sample size calculations for LCGA. However, previous studies analysing latent trajectories in clinical outcomes have successfully identified three to five latent class trajectories through latent class growth analysis in clinical cohorts in primary care with sample sizes of approximately 500 participants (Nicholls et al 2014; Rzewuska et al 2015) and over one thousand participants were included in the LCGA in this study.

As acknowledged in the previous chapter, an *a priori* sample size of 882 was calculated in the protocol for the prospective cohort study using data at all five time-points to detect a smallest meaningful difference in HRQOL of 0.2 standard deviation units detected between two groups (fewer than two gout flares per year and two or more gout flares per year) at a significance of 0.05, power of 90% and autocorrelation of 0.8 (Chandratre et al 2012). As participants with

missing covariate data were omitted from the LMM analysis, the sample size included in each LMM is below the 882 specified *a priori.*

The limitations of using sample size calculations in multilevel modelling has been highlighted due to its reliance on statistical significance and caution has been advised regarding using sample size calculations in multilevel modelling (Twisk, 2006). A particular challenge with calculating sample size and power in multilevel modelling is selecting the fixed-effect which is to be tested (Grace-Martin, 2019) and where associations between several covariates and the outcome are tested this could involve multiple power calculations to determine the chance of detecting a real effect as statistically significant (rejecting the null hypothesis when it is false) (Petrie & Sabin, 2009). The use of post-hoc power analysis has been specifically discouraged in the literature (Hoenig & Heisey, 2001; Zhang et al 2019). As sample size and power calculations are closely related to statistical significance (Twisk, 2006), the interpretation of the width of confidence intervals can indicate how confident we are regarding the relation of the result to the null hypothesis, so it has been proposed that post-hoc power calculations provide no additional insight (Hoenig & Heisey, 2001).

SF-36 PF10 missing data

After the thesis was submitted an error in the original Stata do-file used to calculate the SF-36 PF10 scores was found. This error led to a greater proportion of missing data in the SF-36 PF10 scores used in this thesis. The original analysis has been retained in the thesis but an amended do-file will be used to generate SF-36 PF10 data for reanalysis prior to submission for publication.

9.4 Implications for clinical practice and people living with gout

The findings of this thesis are relevant to clinical practice as the thesis focuses on people living with gout in primary care, where the majority of people living with gout are managed. The

implications of the findings for both clinical practice and people living with gout were discussed with both clinicians and members of the Keele RUG during a PPIE group meeting convened to specifically discuss the results of this study.

When considering the implications of the findings to clinical practice, it is important to reflect on the generalisability of the cohort to the wider population encountered in clinical practice. This was an older cohort who were predominantly male, which is typical of patients seen in primary care (Kuo et al 2015a). Responders in this cohort reported slightly less severe (better treated) gout and slightly better HRQOL compared to non-participants, thus it is important to appreciate that patients in clinical practice may experience more frequent gout flares and worse HRQOL in comparison to the study cohort. Responders to this study are also likely to be less deprived and better educated compared to the wider population.

The plotting of intra-individual gout flare frequency and HRQOL scores over time in this thesis, demonstrates the diverse experiences of people living with gout in primary care. This is an important consideration for management in clinical practice and supports the role of individualised care.

Relevance to clinical guidelines

More members of the infrequent gout flare trajectory class reported using allopurinol and had a supressed urate level, whilst classes with more frequent flares were associated with elevated urate levels. In addition, allopurinol use was associated with an improvement in HRQOL, specifically disease-specific unmet treatment need, however allopurinol use was also associated with lower wellbeing during an attack. More frequent flares were associated with worse disease-specific and generic HRQOL, however HRQOL was also impaired in those experiencing only one flare per year.

It is important to appreciate that these findings are derived from an observational cohort and thus causal interpretations cannot be made. However, these findings are consistent with

existing recommendations in gout management guidelines for people living with gout to strongly consider ULT in those individuals with two or more flares per year, but also to discuss ULT early in the course of the disease and to support patients receiving ULT as reducing urate levels can trigger gout flares (Hui et al 2017).

Both Rheumatologists and GPs have suggested that flare trajectory diagrams could be used as a visual aid to facilitate conversations with patients about management of gout and future outcomes.

Suboptimal gout management

Less than a third of the participants reported taking an allopurinol dose of 300mg or more per day and over 60% of the cohort were missing a serum urate record in their medical records. Such suboptimal management of gout in primary care is consistent with findings in previous studies. Most patients receiving ULT do not receive a dose adjustment to optimise urate levels (Cottrell et al 2013) or are non-adherent to ULT (Scheepers et al 2018). Inadequate serum urate monitoring has also been demonstrated in other cohorts (Wall et al 2010). Due to suboptimal dosing of allopurinol and suboptimal serum urate recording it was not possible to investigate the full potential of the effects of allopurinol or serum urate levels on HRQOL in this cohort. It is possible that the influence may have been greater, if allopurinol use was optimised.

Thus, changes in HRQOL and gout flares were seen in people with suboptimal management within primary care. Consequently, it is perhaps unsurprising that the trajectory of gout flares for the majority of participants did not display any change over time (the majority of participants were allocated to either the infrequent, moderately frequent or frequent and persistent classes) and group-level change in HRQOL scores were small or negligible.

When these results were discussed with the Keele RUG PPIE group they felt the key implications related to the associations between allopurinol use and both flare frequency and HRQOL. PPIE group members also felt that optimising best practice for management of gout was the key implication for clinical practice.

Targeting interventions and resources

The characteristics of participants in the frequent or worsening gout flare trajectories and the characteristics associated with change in HRQOL over three years could have implications for targeting interventions and healthcare resources to people living with gout.

The fact that an improvement in generic HRQOL was associated with living in a less deprived area and more participants in frequent or worsening flare trajectories were classified as most deprived, suggests that outcomes in gout are linked to health inequality. It is thus important that interventions, education, and resources which are targeted at people living with gout are likely to work well in areas of deprivation. Consideration of deprivation when planning healthcare provision would be in line with recommendations in the National Health Service (NHS) long term plan to reduce healthcare inequality in the NHS (National Health Service, 2019).

Attendance at further education was associated with an improvement in disease-specific HRQOL over three years and fewer participants in the frequent and worsening gout flare trajectories had attended further education. Low educational attainment is associated with limited (low or marginal) health literacy (Protheroe et al 2017) suggesting that patient information resources for people with gout should use a range of health information materials which are clear, concise and use plain English (Public Health England, 2015).

Whilst gout is more prevalent in males it is important for clinicians to appreciate the experiences of women living with gout. In this cohort being female was associated with a

deterioration in SF36 PF10 and HAQ-DI over three years. Existing studies have also reported the unique impact of a gout diagnosis on females (Richardson et al 2015) and how being female is a risk factor for non-adherence to ULT (Scheepers et al 2018).

Since deterioration in HRQOL was associated with the presence of frequent flares, oligo/polyarticular flares, body pain, worse pain, more severe depression, and anxiety, and a higher BMI, obtaining information about these factors could help to identify people who should be offered ULT.

Impact of findings for people with gout and the general public

The Keele PPIE RUG members felt that dissemination of these findings could have an important role in increasing awareness of gout in the general population. The group suggested that highlighting that gout can be associated with frequent painful inflammatory flares which lead to impairment of HRQOL might help to increase recognition and awareness that gout is as significant chronic inflammatory arthritis, rather than a humorous self-inflicted condition.

9.5 Implications for future research

Replication in other cohorts

As this is the first time that gout flare trajectories have been identified, it would be prudent to investigate gout flare trajectories in other settings and populations. Future analysis of flare trajectories could be undertaken in inception cohorts to gain a picture of the natural history of gout flares from disease onset. A range of different covariates and outcomes were included thus it is possible that some of the associations observed may only exist in this cohort or may have arisen due to chance. Whilst there is consistency with the factors associated with change in HRQOL and the factors associated with HRQOL in people living with gout in previous cross-sectional studies, these should be explored in other cohorts. In addition, as three years may

not be long enough to capture the full extent of change in a chronic illness such as gout, it may also be beneficial to investigate change in HRQOL over a longer period.

Alternative methods of analysis of change over time

SF-36 PF10 scores displayed negligible change at a group level over time and the amount of variation over time in SF-36 PF10 scores did not warrant a random slope at a group level. Little change was also observed in SF-36 PF10 scores at a group level after adjustment for gout-specific, comorbid, socio-demographic and other factors in the LMM. However, plots of intra-individual change in SF36 PF10 did show some variation in change in HRQOL scores over time for some individuals. Thus, an alternative method of statistical modelling, such as latent class growth analysis, could be used in the future to identify individuals with different trajectories of generic HRQOL scores over time and describe the characteristics of participants with different trajectories.

Studies to address modifiable factors

Several modifiable factors, including depression, obesity and pain, were associated with deterioration in disease-specific and/or generic HRQOL over three years. Primary care trials of intervention for obesity, depression, and/ or pain could investigate whether deterioration in HRQOL in people living with gout could be attenuated through identification and intervention. A previous RCT which provided an intervention of depression care management to older persons living with arthritis reported less interference with daily activities and lower pain intensity (Lin et al 2003). A RCT of optimised anti-depressant therapy followed by a pain management programme as the intervention, reported lower depression severity scores, a reduction in pain severity and disability, in a US study of primary care patients with musculoskeletal conditions (Kroenke et al 2009). A systematic review by Nielsen et al (2017) identified 10 longitudinal weight loss studies in people living with gout and reported an effect of weight loss to reduce gout flares. However, the lack of measurement of other outcomes

such as HRQOL and pain in these studies was noted. Importantly studies included were low quality, particularly regarding bias due to confounding and departures from intended interventions, and rigorous RCTs are needed (Nielsen et al 2017).

Prognosis research

Due to the paucity of existing research the statistical analyses undertaken were intentionally exploratory in nature. Future prospective research relating to gout flares and HRQOL in people living with gout could follow a prognosis research approach, where the relationship between future outcomes and baseline factors are investigated (Hemingway et al 2013). This could include prognostic factor research, where a factor's prognostic ability would be examined across multiple studies thus replicating and confirming the factor's prognostic ability (Riley et al 2013). Prognostic research might then be extended further to develop prognostic models which combine multiple predictors to calculate the risk of the outcome for individual patients (Steyerberg et al 2013). Whilst such prognostic models can assist clinicians with their prediction of a patient's future outcome and enhance informed decision making, they require careful development, validation and evaluation in clinical practice (Steyerberg et al 2013).

9.6 Conclusion

Until now little has been known about change in gout flares and HRQOL over time in people living with gout, and the factors associated with worse outcomes. This thesis investigated the change in gout flares, disease-specific HRQOL (GIS subscales) and generic HRQOL (SF-36 PF10 and HAQ-DI) over three years in a prospective cohort study of people living with gout in primary care. Six distinct gout flare trajectory classes were identified. The infrequent flare class had the lowest mean serum urate level and the highest proportion of patients taking allopurinol, whilst frequent flare classes had more patients who were socioeconomically

deprived, obese, had chronic kidney disease, and a history of oligo/polyarticular flares. Factors associated with deterioration in both disease-specific and generic HRQOL included more frequent gout flares, a history of oligo/polyarticular flares, allopurinol use, the presence of body pain, worse pain severity, and worse depression score. Factors associated with improvement in disease-specific HRQOL included allopurinol use, a longer disease duration, older age, and attendance at further education. Factors associated with improvement in generic HRQOL included living in a less socioeconomically deprived area and consuming alcohol more frequently.

These thesis findings are relevant to clinical practice as they investigate the outcomes of people living with gout in primary care, where the majority of people are managed. Gout is a common inflammatory arthritis with a well understood pathophysiology and readily available effective treatment, yet a significant number of patients continue to experience poor outcomes. These observational findings are consistent with existing clinical guidelines that advise that the option of ULT should be discussed with all people living with gout but particularly people experiencing two or more flares and oligo/polyarticular flares (Hui et al 2017). The identification of people at risk of worse outcomes, for example worse pain, depression or obesity, could help to target interventions and healthcare resources for people living with gout. Interventions, education and resources targeted at people living with gout need to be able to work well in areas of deprivation and in individuals with lower educational attainment.

10 References

Abhishek, A., Roddy, E. & Doherty, M. (2017) Gout a guide for the general and acute physicians. *Clinical Medicine*. 17(1), 54–59. doi:10.7861/clinmedicine.17-1-54

Abhishek, A., Valdes, A. M., Zhang, W. & Doherty, M. (2016) Association of Serum Uric Acid and Disease Duration With Frequent Gout Attacks: A Case–Control Study. *Arthritis Care & Research*. 68(10), 1573–1577. doi: 10.1002/acr.22855

Akaike, H. (1987) Factor analysis and AIC. *Psychometrika. 52*(3), 317–332. doi.org/10.1007/BF02294359

Alvarez-Hernández, E., Peláez-Ballestas, I., Vázquez-Mellado, J., Terán-Estrada, L., Bernard-Medina, A. G., Espinoza, J., Aceves-Avila, F. J., Goycochea-Robles, M. V, Garza, M., Ventura, L. & Burgos-Vargas, R. (2008) Validation of the Health Assessment Questionnaire disability index in patients with gout. *Arthritis and Rheumatism*. 59(5), 665–669. doi: 10.1002/art.23575

Alvarez-Hernández, E., Zamudio-Lerma, J. A., Burgos-Martínez, G., Álvarez-Etchegaray, S. E., Peláez-Ballestas, I. & Vázquez-Mellado, J. (2009) Measurement of health-related quality of life and functional capacity in patients with chronic tophaceous gout. *Reumatología Clínica* (English Edition). 5(3), 103–108. doi:10.1016/S2173-5743(09)70101-7

Alvarez-Nemegyei, J., Cen-Piste, J. C., Medina-Escobedo, M. & Villanueva-Jorge, S. (2005) Factors associated with musculoskeletal disability and chronic renal failure in clinically diagnosed primary gout. *The Journal of Rheumatology*. *32*(10), 1923–1927.

Andruff, H., Carraro, N., Thompson, A., Gaudreau, P. & Louvet, B. (2009) Latent class growth modelling: A tutorial. *Tutorials in Quantitative Methods for Psychology*. *5*(1), 11–24. doi:10.20982/tqmp.05.1.p011

Annemans, L., Spaepen, E., Gaskin, M., Bonnemaire, M., Malier, V., Gilbert, T. & Nuki, G. (2008) Gout in the UK and Germany: Prevalence, comorbidities and management in general practice 2000-2005. *Annals of the Rheumatic Diseases*. 67(7), 960–966. doi: 10.1136/ard.2007.076232

Bakas, T., McLennon, S. M., Carpenter, J. S., Buelow, J. M., Otte, J. L., Hanna, K. M., Ellett, M. L., Hadler, K. A. & Welch, J. L. (2012) Systematic review of health-related quality of life models. *Health & Quality of Life Outcomes*. 10(1), 134–145. doi: 10.1186/1477-7525-10-134

Baker, D., Mead, N. & Campbell, S. (2002) Inequalities in morbidity and consulting behaviour for socially vulnerable groups. *British Journal of General Practice*. 52(475), 124–130.

Bankhead, C., Aronson, J. & Nunan, D. (2018) Catalogue of bias: attrition bias. *BMJ Evidence-Based Medicine*. 23 (1), 21–22. doi:10.1136/ebmed-2017-110883.

Barclay-Goddard, R., Lix, L. M., Tate, R., Weinberg, L. & Mayo, N. E. (2011) Health-related quality of life after stroke: Does response shift occur in self-perceived physical function? *Archives of Physical Medicine and Rehabilitation*. 92(11), 1762–1769. doi:10.1016/j.apmr.2011.06.013

Barnett, A. G., van der Pols, J. C. & Dobson, A. J. (2005) Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*. 34(1), 215–220. doi:10.1093/ije/dyh299

Bauer, D. J. & Curran, P. J. (2003) Distributional Assumptions of Growth Mixture Models: Implications for Overextraction of Latent Trajectory Classes. *Psychological Methods. 8*(3), 338–363. doi: 10.1037/1082-989X.8.3.338

Becker, M. A., Schumacher, H. R., Benjamin, K. L., Gorevic, P., Greenwald, M., Fessel, J., Edwards, L., Kawata, A. K., Frank, L., Waltrip, R., Maroli, A., Huang, B., Sundy, J. S., Weisman, M., Reginato, A., Griffin, R., Goldberger, E., Simkin, P., Morgan, J., Kivitz, A., Weitz, M., Fung, H. S. E., Fudman, E., Cronin, M., Peller, J. S., Lies, R., Block, J., Rosenthal, P., Merrill, J. & Dewees, G. (2009) Quality of life and disability in patients with treatment-failure gout. *Journal of Rheumatology*. 36(5), 1041–1048. doi:10.3899/jrheum.071229

Bennett, D.A. (2001) How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health*. 25 (5), 464–469. doi:10.1111/j.1467-842X.2001.tb00294.x

Bentler, P. M. (1990) Comparative fit indexes in structural models. *Psychological Bulletin*. 107(2), 238-246.

Bentler, P. M. & Bonett, D. G. (1980) Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*. *88*(3), 588–606. doi:10.1037/0033-2909.88.3.588

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J. & Crotty, K. (2011) Low health literacy and health outcomes: an updated systematic review. *Annals of Internal Medicine.* 155(2), 97–107. doi:10.7326/0003-4819-155-2-201107190-00005

Berlin, K. S., Parra, G. R. & Williams, N. A. (2014) An introduction to latent variable mixture modeling (part 2): longitudinal latent class growth analysis and growth mixture models. *Journal of Pediatric Psychology. 39*(2), 188–203. doi:10.1093/jpepsy/jst085

Berrington, A., Smith, W.F. & Sturgis, P. (2006) An Overview of Methods for the Analysis of Panel Data. *Economic and Social Research Council (ESRC) National Centre for Research Methods Briefing Paper*, (November), 1–57.

Biering, K., Hjollund, N.H. & Frydenberg, M. (2015) Using multiple imputation to deal with missing data and attrition in longitudinal studies with repeated measures of patient-reported outcomes. *Clinical Epidemiology*. 7, 91–106. doi:10.2147/CLEP.S72247

Blagojevic-Bucknall, M., Mallen, C., Muller, S., Hayward, R., West, S., Choi, H. & Roddy, E. (2019) The Risk of Gout Among Patients With Sleep Apnea: A Matched Cohort Study. *Arthritis and Rheumatology*. 71(1), 154–160. doi: 10.1002/art.40662

Bland, M. (2015) *An introduction to medical statistics*. 4th edition. Oxford, Oxford University Press.

Blome, C. & Augustin, M. (2015) Measuring change in quality of life: Bias in prospective and retrospective evaluation. *Value in Health*. 18(1), 110–115. doi:10.1016/j.jval.2014.10.007

Bohning, D., Dietz, E., Schlattmann, P., Mendonca, L., Mendonga, L. & Kirchner, U. (1999) The Zero-Inflated Poisson Model and the Decayed, Missing and Filled Teeth Index in Dental Epidemiology. *Journal of the Royal Statistical Society. 162*(2), 195–209. doi: 10.1111/1467-985X.00130

Bowen-Davies, Z., Muller, S., Mallen, C. D., Hayward, R. A. & Roddy, E. (2018) Gout Severity, Socioeconomic Status, and Work Absence: A Cross-Sectional Study in Primary Care. *Arthritis Care & Research*. 70(12), 1822–1828. doi: 10.1002/acr.23562

Bowling, A. (2003) Current state of the art in quality of life measurement. In: Carr, A. J., Higginson, I. J. & Robinson, P. (eds.) *Quality of life*. London, BMJ Books.

Bowling, A. (2005) *Measuring health A Review of quality of life measurement scales*. 3rd edition. Maidenhead, Open University Press.

Bowling, A. (2009) *Research Methods in Health Investigating Health and Health Services*. 3rd edition. Maidenhead, Open University Press.

Bowling, A. (2014). *Research Methods in Health Investigating Health and Health services*. 4th edition. Maidenhead, Open University Press.

Bowling, A., Bond, M., Jenkinson, C. & Lamping, D. L. (1999) Short Form 36 (SF-36) Health Survey Questionnaire: Which normative data should be used? Comparisons between the norms provided by the Omnibus Survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *Journal of Public Health Medicine*. *21*(3), 255–270. doi:10.1093/pubmed/21.3.255

Boynton, P. M. & Greenhalgh, T. (2004) Selecting, designing, and developing your questionnaire. *British Medical Journal*. *328*, 1312–1315. doi:10.1136/bmj.328.7451.1312

Brauer, G. W. and Prior, I. A. (1978) A prospective study of gout in New Zealand Maoris. *Annals of the rheumatic diseases*. 37(5), 466–72. doi: 10.1136/ard.37.5.466

Brazier, J. E., Harper, R., Jones, N. M., O'Cathain, A., Thomas, K. J., Usherwood, T. & Westlake, L. (1992) Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *British Medical Journal. 305*, 160–164. doi:10.1136/bmj.305.6846.160

Brenes, G. A. (2007) Anxiety, depression, and quality of life in primary care patients. *Primary Care Companion to the Journal of Clinical Psychiatry*. 9(6), 437–443. doi: 10.4088/PCC.v09n0606

Brilleman, S.L., Pachana, N.A. & Dobson, A.J. (2010) The impact of attrition on the representativeness of cohort studies of older people. *BMC Medical Research Methodology*. 10, 71. doi:10.1186/1471-2288-10-71

Bruce, B. & Fries, J.F. (2003) The Stanford health assessment questionnaire HAQ A review of its history issues progress and documentation. *The Journal of Rheumatology*. 30, 167–178.

Bruce, N., Pope, D. & Stanistreet, D. (2008) *Quantitative Methods for Health Research*. Chichester, John Wiley & Sons.

Bruderer, S., Bodmer, M., Jick, S. S. & Meier, C. R. (2014) Use of diuretics and risk of incident gout: A population-based case-control study. *Arthritis and Rheumatology*. 66(1), 185–196. doi: 10.1002/art.38203

Bruderer, S. G., Meier, C. R., Jick, S. S. & Bodmer, M. (2017) The association between thyroid disorders and incident gout: Population-based case–control study. *Clinical Epidemiology*. 9, 205–215. doi: 10.2147/CLEP.S128627

Burant, C. J. (2016) Latent Growth Curve Models: Tracking changes over time. *The International Journal of Aging and Human Development*. *82*(4), 336–350. doi: 10.1177/0091415016641692

Bursill, D., Taylor, W. J., Terkeltaub, R., Kuwabara, M., Merriman, T. R., Grainger, R., Pineda, C., Louthrenoo, W., Edwards, N.L., Andres, M., Vargas-Santos, A.B., Roddy, E., Pacsart, T., Lin, C., Perex-Ruis, F., Tedeschi, S. K., Kim, S.C., Harrold, L.R., McCArthy, G., Kumar, N., Chapman, P.T., Tausche, A., Vazquez-Mellado, J., Gtierrez, M., Casterlar-Pinheiro, G., Richette, P.,

Pascual, E., Fisher, M.C., Burgos-Vargas, R., Robinson, P.C., Singh, J.A., Jansen, T.L., Saag, K.G., Slot, O., Uhlig, T., Soloman, D.H., Keenan, R.T., Scire, C. A., Biernat-Kaluza, E., Dehlin, M., Nuki, G., Schlesinger, N., Janssen, M., Stamp, L.K., Sivera, F., Reginato, A.M., Jacobsson, L., Liote, F., Ea, H., Rosenthal, A., Bardin, T., Choi, H.K., Hershfield, M.S., Czegley, C. Choi, S.J. & Dalbeth, N. (2019) Gout, Hyperuricemia, and Crystal-Associated Disease Network Consensus Statement Regarding Labels and Definitions for Disease Elements in Gout. *Arthritis Care & Research*. 71(3), 427–434. doi:10.1002/acr.23607

Calderon, J. L., Morales, L. S., Liu, H. & Hays, R. D. (2006) Variation in the readability of items within surveys. *American Journal of Medical Quality*. *21*(1), 49–56. doi:10.1177/1062860605283572

Calvert, M. J. & Freemantle, N. (2003) Use of health-related quality of life in prescribing research. Part 1: Why evaluate health-related quality of life? *Journal of Clinical Pharmacy and Therapeutics.* 28(6), 513–521.

Campion, E.W., Glynn, R.G. & Delabry, L. O. (1987) Asymptomatic hyperuricemia. Risks and consequences in the normative aging study. *The American Journal of Medicine*. 82(3), 421–426. doi: 10.1016/0002-9343(87)90441-4

Carr, A., Gibson, B. & Robinson, P.G. (2003) Is quality of life determined by expectations or experiences. In: Carr, A., Higginson, I. J. & Robinson, P. G. (eds.) *Quality of Life*. London, BMJ Books. 9-18.

Cea Soriano, L., Rothenbacher, D., Choi, H. K. & García Rodríguez, L. A. (2011) Contemporary epidemiology of gout in the UK general population. *Arthritis Research & Therapy.* 13, R39. doi:10.1186/ar3272

Chandratre, P., Mallen, C.D., Richardson, J., Muller, S., Hider, S., Rome, K., Blagojevic-Bucknall, M. & Roddy, E. (2018) Health-related quality of life in gout in primary care: Baseline findings from a cohort study. *Seminars in Arthritis and Rheumatism*. 48(1), 61–69. doi:10.1016/j.semarthrit.2017.12.005

Chandratre, P., Mallen, C.D., Richardson, J., Rome, K., Bailey, J., Gill, R., Hiders, S., Mason, J., Mayson, Z., Muller, S., Purcell, C., Titley, J., Wathall, S., Zwierska, I. & Roddy, E. (2012) Prospective observational cohort study of Health Related Quality of Life (HRQOL), chronic foot problems and their determinants in gout: a research protocol. *BMC Musculoskeletal Disorders*. *13*, 219. doi:10.1186/1471-2474-13-219

Chandratre, P., Mallen, C. D., Roddy, E., Liddle, J. & Richardson, J. (2016) "You want to get on with the rest of your life": a qualitative study of health-related quality of life in gout. *Clinical Rheumatology*. 35(5), 1197–1205. doi:10.1007/s10067-015-3039-2

Chandratre, P., Roddy, E., Clarson, L., Richardson, J., Hider, S. L. & Mallen, C. D. (2013) Health-related quality of life in gout: a systematic review. *Rheumatology*. *52*(11), 2031–2040. doi:10.1093/rheumatology/ket265

Chatfield, M.D., Brayne, C.E. & Matthews, F.E. (2005) A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. *Journal of Clinical Epidemiology*. 58(1), 13–19. doi: 10.1016/j.jclinepi.2004.05.006

Cheng, J., Edwards, L.J., Maldonado-Molina, M.M., Komro, K.A. & Muller, K.E. (2010) Real Longitudinal Data Analysis for Real People: Building a Good Enough Mixed Model. *Statistics in medicine.* 29(4), 504–520. doi:0.1002/sim.3775.Real

Cheong, Y. F., Fotiu, R. P. & Raudenbush, S. W. (2001) Efficiency and Robustness of Alternative Estimators for Two- and Three-level Models: The Case of NAEP. *Journal of Educational and Behavioural Statistics*. 26(4), 411–429. doi:10.3102/10769986026004411

Chhana, A. & Dalbeth, N. (2015) The Gouty Tophus: a Review. *Current Rheumatology Reports*. 17(19), 1–9. doi: 10.1007/s11926-014-0492-x

Chinchilla, S. P., Doherty, M. & Abhishek, A. (2019) Gout Activity Score has predictive validity and is sensitive to change: results from the Nottingham Gout Treatment Trial (Phase II). *Rheumatology*. 58(8), 1378–1382. doi: 10.1093/rheumatology/key446

Chisholm, J. (1990) The Read clinical classification. *British Medical Journal*. *300*, 1092. doi:10.1136/bmj.300.6732.1092

Choi, H.K., Atkinson, K., Karlson, E.W., Willett, W. & Curhan, G. (2004a) Alcohol intake and risk of incident gout in men: A prospective study. *The Lancet*. 363 (9417), 1277–1281. doi:10.1016/S0140-6736(04)16000-5

Choi, H.K., Atkinson, K., Karlson, E.W., Willett, W. & Curhan, G. (2004b) Purine-Rich Foods, Dairy and Protein Intake, and the Risk of Gout in Men. *The New England Journal of Medicine*. 350 (11), 1093–1103. doi:10.1056/NEJMoa035700

Choi, H. K. & Curhan, G. (2008) Soft drinks, fructose consumption, and the risk of gout in men: Prospective cohort study. *British Medical Journal*. 336(7639), 309–312. doi: 10.1136/bmj.39449.819271.BE

Choi, H. K., Ford, E. S., Li, C. & Curhan, G. (2007) Prevalence of the metabolic syndrome in patients with gout: The Third National Health and Nutrition Examination Survey. *Arthritis Care & Research*. 57(1), 109–115. doi:10.1002/art.22466

Choi, H.K., Gao, X. & Curhan, G. (2009) Vitamin C intake and the risk of gout in men: a prospective study. *Archives of internal medicine*. 169 (5), 502–507. doi:10.1001/archinternmed.2008.606.Vitamin

Choi, H. K., Mount, D. B. & Reginato, A. M. (2005) Pathogenesis of gout. *Annals of Internal Medicine.* 143(7), 499–516. doi: 10.7326/0003-4819-143-7-200510040-00009

Choi, H. K., Soriano, L. C., Zhang, Y. & Garciá Rodriǵuez, L. A. (2012) Antihypertensive drugs and risk of incident gout among patients with hypertension: population-based case-control study. *British Medical Journal*. 344, d8190 doi: 10.1136/bmj.d8190

Choi, H. K., Willett, W. & Curhan, G. (2007) Coffee consumption and risk of incident gout in men a prospective study. *Arthritis and Rheumatism*. 56(6), 2049–2055. doi: 10.1002/art.22712

Cieza, A. & Stucki, G. (2008) The international classification of functioning disability and health: Its development process and content validity. *European Journal of Physical and Rehabilitation Medicine*. 44(3), 303–313. doi: 18762740

Clark, S. & Muthén, B. O. (2009) Relating Latent Class Analysis Results to Variables not Included in the Analysis. StatisticalInnovations.Com, 1–55. Retrieved from https://www.statmodel.com/download/relatinglca.pdf Accessed: September 2017

Clarson, L. E., Hider, S. L., Belcher, J., Heneghan, C., Roddy, E. & Mallen, C. D. (2014) Increased risk of vascular disease associated with gout: a retrospective, matched cohort study in the UK clinical practice research datalink. *Annals of the Rheumatic Diseases*. 74(4), 642–7. doi: 10.1136/annrheumdis-2014-205252

Cnaan, A., Laird, N. M. & Slasor, P. (1997) Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Statistics in Medicine. 16*(20), 2349–2380. doi:10.1002/(SICI)1097-0258(19971030)16:20<2349::AID-SIM667>3.0.CO;2-E

Cohen, J. (1992) A Power Primer. *Psychological Bulletin*. 112(1), 155–159.

Colwell, H. H., Hunt, B. J., Pasta, D. J., Palo, W. A., Mathias, S. D. & Joseph-Ridge, N. (2006) Gout Assessment Questionnaire: Initial results of reliability, validity and responsiveness. *International Journal of Clinical Practice. 60*(10), 1210–1217. doi:10.1111/j.1742-1241.2006.01104.x

Cottrell, E., Chambers, R. & O'Connell, P. (2012) Using simple telehealth in primary care to reduce blood pressure: A service evaluation. *BMJ Open*. 2, e001391. doi:10.1136/bmjopen-2012-001391.

Cottrell, E., Crabtree, V., Edwards, J. J. & Roddy, E. (2013) Improvement in the management of gout is vital and overdue: An audit from a UK primary care medical practice. BMC Family Practice. *BMC Family Practice*. 14, 170. doi:10.1186/1471-2296-14-170

Cudjoe, T. K. M., Roth, D. L., Szanton, S. L., Wolff, J. L., Boyd, C. M. & Thorpe, R. J. (2020) The Epidemiology of Social Isolation: National Health & Aging Trends Study. *The Journals of Gerontology Series B*. 75(1), 107-113. doi:10.1093/geronb/gby037

Cumming, J. & Goldstein, H. (2016) Handling attrition and non-response in longitudinal data with an application to a study of Australian youth. *Longitudinal and Life Course Studies*. 7(1), 53–63. doi: 10.14301/llcs.v7i1.342

Curran, P. J. & Hussong, A. M. (2003) The Use of Latent Trajectory Models in Psychopathology Research. *Journal of Abnormal Psychology*. *112*(4), 526–544. doi:10.1037/0021-843X.112.4.526

Curran, P. J., Obeidat, K. & Losardo, D. (2010) Twelve frequently asked questions about growth curve modeling. *Journal of Cognitive Development*. *11*(2), 121–136. doi:10.1080/15248371003699969

Dagne, G. A. & Huang, Y. (2012) Mixed-effects Tobit joint models for longitudinal data with skewness, detection limits, and measurement errors. *Journal of Probability and Statistics*. 614102. doi:10.1155/2012/614102

Dalbeth, N., Ames, R., Gamble, G. D., Horne, A., Wong, S., Kuhn-Sherlock, B., MacGibbon, A., McQueen, F. M., Reid, I. R. & Palmano, K. (2012) Effects of skim milk powder enriched with glycomacropeptide and G600 milk fat extract on frequency of gout flares: A proof-of-concept randomised controlled trial. *Annals of the Rheumatic Diseases*. 71(6), 929–934. doi: 10.1136/annrheumdis-2011-200156

Dalbeth, N., House, M. E., Horne, A., Te Karu, L., Petrie, K. J., McQueen, F. M., & Taylor, W. J. (2013) The experience and impact of gout in Māori and Pacific people: a prospective observational study. *Clinical Rheumatology*. *32*(2), 247–251. doi:10.1007/s10067-012-2110-5

Dalbeth, N., Lauterio, T. J., & Wolfe, H. R. (2014) Mechanism of action of colchicine in the treatment of gout. *Clinical Therapeutics*. 36(10), 1465–1479. doi:10.1016/j.clinthera.2014.07.017

Dalbeth, N., Merriman, T. R. & Stamp, L. K. (2016) Gout. *The Lancet*. 388(10055), 2039-2052. doi:10.1016/S0140-6736(16)00346-9

Dalbeth, N., Petrie, K. J., House, M., Chong, J., Leung, W., Chegudi, R., Horne, A., Gamble, G., McQueen, F. M. & Taylor, W. J. (2011) Illness perceptions in patients with gout and the relationship with progression of musculoskeletal disability. *Arthritis Care & Research*. 63(11), 1605–1612. doi:10.1002/acr.20570

Dalbeth, N., Phipps-Green, A., Frampton, C., Neogi, T., Taylor, W.J. & Merriman, T.R. (2018) Relationship between serum urate concentration and clinically evident incident gout: an individual participant data analysis. *Annals of the Rheumatic Disease*. 77(7), 1048–1052. doi: 10.1136/annrheumdis-2017-212288

Dalbeth, N., Pool, B., Gamble, G. D., Smith, T., Callon, K. E., McQueen, F. M. & Cornish, J. (2010) Cellular characterization of the gouty tophus: A quantitative analysis. *Arthritis and Rheumatism*. 62(5), 1549–1556. doi:10.1002/art.27356

Dalbeth, N., Stamp, L. K. & Merriman, T. R. (2017) The genetics of gout: Towards personalised medicine? *BMC Medicine*. 15,108. doi:10.1186/s12916-017-0878-5

de Lautour, H., Dalbeth, N. & Taylor, W. J. (2015) Outcome Measures for Gout Clinical Trials: a Summary of Progress. *Current Treatment Options in Rheumatology*. 1(2), 156–166. doi: 10.1007/s40674-015-0014-7

Deeg, D.J.H. (2002) Attrition in longitudinal population studies: Does it affect the generalizability of the findings?: An introduction to the series. *Journal of Clinical Epidemiology*. 55(3), 213–215. doi:10.1016/S0895-4356(01)00472-3

Del Boca, F. K. & Darkes, J. (2003) The validity of self-reports of alcohol consumption: state of the science and challenges for research. *Addiction.* 98(s2), 1–12. doi:10.1046/j.1359-6357.2003.00586.x

Department for Communities and Local Government (2011) The English Indices of Deprivation 2010. Available from: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2010 Accessed: March 2018

Detmar, S. B., Muller, M. J., Schornagel, J. H., Wever, L. D. V. & Aaronson, N. K. (2002) Health-Related Quality-of-Life Assessments and Patient-Physician Communication. *Journal of the American Medical Association*. 288(23), 3027-3034. doi:10.1001/jama.288.23.3027

Diggle, P.J., Liang, K.Y. & Zeger, S. (1994). *Analysis of Longitudinal Data*. Oxford, Oxford University Press.

Dijkers, M. (2007) '"What's in a name?" The indiscriminate use of the "Quality of life" label, and the need to bring about clarity in conceptualizations. *International Journal of Nursing Studies.* 44(1), 153–155. doi:10.1016/j.ijnurstu.2006.07.016

Doherty, M., Jenkins, W., Richardson, H., Sarmanova, A., Abhishek, A., Ashton, D., Barclay, C., Doherty, S., Duley, L., Hatton, R., Rees, F., Stevenson, M. & Zhang, W. (2018) Efficacy and cost-effectiveness of nurse-led care involving education and engagement of patients and a treat-to-target urate-lowering strategy versus usual care for gout: a randomised controlled trial. *The Lancet*. 392(10156), 1403–1412. doi:10.1016/S0140-6736(18)32158-5

Domecq, J. P., Prutsky, G., Elraiyah, T., Wang, Z., Nabhan, M., Shippee, N., Brio, J.P., Boehmer, K., Hasan, R., Firwana, B., Erwin, P., Eton, D., Sloan, J., Montori, V., Asi, N., Abu Darbh, A. M., & Murad, M. H. (2014) Patient engagement in research: a systematic review. *BMC Health Services Research*. 14(1), 89. doi:10.1186/1472-6963-14-89

Dong, Y. & Peng, C.Y.J. (2013) Principled missing data methods for researchers. *SpringerPlus*. 2, 222. doi:10.1186/2193-1801-2-222

Drieling, R. L., LaCroix, A. Z., Beresford, S. A. A., Boudreau, D. M., Kooperberg, C. & Heckbert, S. R. (2016) Validity of Self-Reported Medication Use Compared with Pharmacy Records in a Cohort of Older Women: Findings from the Women's Health Initiative. *American Journal of Epidemiology*. *184*(3), 233–238. doi:10.1093/aje/kwv446

Dufour, A. B., Broe, K. E., Nguyen, U. S. D. T., Gagnon, D. R., Hillstrom, H. J., Walker, A. H., Kivell, E. & Hannan, M. T. (2009) Foot pain: Is current or past shoewear a factor? *Arthritis Care & Research*. 61 (10), 1352-1358. doi:10.1002/art.24733

Dumville, J.C., Torgerson, D.J. & Hewitt, C.E. (2006) Reporting attrition in randomised controlled trials. *British Medical Journal*. 332 (7547), 969–971. doi:10.1136/bmj.332.7547.969

Duncan, T.E. & Duncan, S. C. (2009) The ABC's of LGM: An introductory guide to latent variable growth curve modeling. *Social and Personality Psychology Compass*. *3*(6), 979–991. doi:10.1111/j.1751-9004.2009.00224.x

Dunn, K. M., Jordan, K. & Croft, P. R. (2003) Does questionnaire structure influence response in postal surveys? *Journal of Clinical Epidemiology*. 56(1), 10–16. doi:10.1016/S0895-4356(02)00567-X

Duyck, S. D., Petrie, K. J. & Dalbeth, N. (2016) '"You Don' t Have to Be a Drinker to Get Gout, But It Helps": A Content Analysis of the Depiction of Gout in Popular Newspapers. *Arthritis Care & Research*. 68(11), 1721–1725. doi: 10.1002/acr.22879

Edgerton, J.D., Roberts, L.W., & von Below, S. (2012) Education and Quality of Life. In: Land, K. C., Michalos, A. C. & Sirgy, M. J. (eds.). *Handbook of Social Indicators and Quality of Life Research.* 265–296.

Edwards, N.L, Singh, J.A., Troum, O., Yeo, A.E. & Lipsky, P.E. (2019) Characterization of patients with chronic refractory gout who do and do not have clinically apparent tophi and their response to pegloticase. *Rheumatology*. 58 (8), 1422–1431. doi:10.1093/rheumatology/kez017

Edwards, P. J., Roberts, I., Clarke, M. J., DiGuiseppi, C., Wentz, R., Kwan, I., Cooper, R., Felix, L.M. & Pratap, S. (2009) Methods to increase response to postal and electronic questionnaires. *Cochrane Database of Systematic Reviews*. 3, MR000008. doi:10.1002/14651858.MR000008.pub4

Eisele, M., Kaduszkiewicz, H., König, H. H., Lange, C., Wiese, B., Prokein, J., Weyerer, S., Werle, J., Riedel-Heller, S.G., Luppa, M., Heser, K., Koppara, A., Mosh, E., Weeg, D., Fuchs, A., Pentzek, M., Maier, W.Scherer, M., & Hajek, A. (2015) Determinants of health-related quality of life in older primary care patients: Results of the longitudinal observational AgeCoDe Study. *British Journal of General Practice*. 65(640), e716–e723. doi:10.3399/bjgp15X687337

Enomoto, A., Kimura, H., Chairoungdua, A., Shigeta, Y., Jutabha, P., Ho Cha, S., Hosoymada, M., Takeda, M., Sekine, T., Igarashi, T., Matsuo, H., Kikuchi, Y., Oda, T., Ichida, K., Hosoya, T., Shimokata, K., Niwa, T., Kanai, Y. & Endou, H. (2002) Molecular identification of a renal urate anion exchanger that regulates blood urate levels. *Nature*. 417(6887), 447–452. doi:10.1038/nature742

Evans, P. L., Prior, J. A., Belcher, J., Mallen, C. D., Hay, C. A. & Roddy, E. (2018) Obesity, hypertension and diuretic use as risk factors for incident gout: A systematic review and meta-analysis of cohort studies. *Arthritis Research & Therapy*. 20, 136. doi:10.1186/s13075-018-1612-1

Fayers, P.M. & Machin, D. (2016) *Quality of Life; The assessment, analysis and reporting of patient-reported outcomes*. Chichester, John Wiley & Sons.

Felce, D. & Perry, J. (1995) Quality of life: its definition and measurement. *Research Developmental Disabilities*. 16(1), 51–74. doi:10.1016/0891-4222(94)00028-8

Feldman, B. J., Masyn, K. E. & Conger, R. D. (2009) New Approaches to Studying Problem Behaviors: A Comparison of Methods for Modeling Longitudinal, Categorical Adolescent Drinking Data. *Developmental Psychology*. *45*(3), 652–676. doi:10.1037/a0014851

Fewtrell, M.S., Kennedy, K., Singhal, A., Martin, R.M., Ness, A., Hadders-Algra, M., Koletzko, B. & Lucas, A. (2008) How much loss to follow-up is acceptable in long-term randomised trials and prospective studies? *Archives of Disease in Childhood*. 93 (6), 458–461. doi:10.1136/adc.2007.127316

Fielding, S., Fayers, P.M. & Ramsay, C.R. (2009) Investigating the missing data mechanism in quality of life outcomes: A comparison of approaches. *Health and Quality of Life Outcomes*. 7, 57. doi:10.1186/1477-7525-7-57

Fincham, J.E. (2008) Response rates and responsiveness for surveys, standards, and the Journal. *American Journal of Pharmaceutical Education*. 72, 43. doi:10.5688/aj720243

Fitzpatrick, R., Fletcher, A., Gore, S., Jones, D., Spiegelhalter, D. & Cox, D. (1992) Quality of life measures in health care. I: Applications and issues in assessment. *British Medical Journal*. 305(6861), 1074–7. doi:10.1136/bmj.305.6861.1074

Fu, T., Cao, H., Yin, R., Zhang, L., Zhang, Q., Li, L. & Gu, Z. (2017) Associated factors with functional disability and health-related quality of life in Chinese patients with gout: A case-control study. *BMC Musculoskeletal Disorders*. 18, 429. doi: 10.1186/s12891-017-1787-7

Gaffo, A. L., Dalbeth, N., Saag, K. G., Singh, J. A., Rahn, E. J., Mudano, A. S., Chen, Y. H., Lin, C. T., Bourke, S., Louthrenoo, W., Vazquez-Mellado, J., Hernández-Llinas, H., Neogi, T., Vargas-Santos, A. B., da Rocha Castelar-Pinheiro, G., Amorim, R. B. C., Uhlig, T., Hammer, H. B., Eliseev, M., Perez-Ruiz, F., Cavagna, L., McCarthy, G. M., Stamp, L. K., Gerritsen, M., Fana, V., Sivera, F. & Taylor, W. (2018) Brief Report: Validation of a Definition of Flare in Patients With Established Gout. *Arthritis and Rheumatology*. 70(3), 462–467. doi:10.1002/art.40381

Gaffo, A. L., Schumacher, H. R., Saag, K. G., Taylor, W. J., Dinnella, J., Outman, R., Chen, L., Dalbeth, N., Sivera, F., Vazquez-Mellado, J., Chou, C. T., Zeng, X., Perez-Ruiz, F., Kowalski, S. C., Goldenstein-Schainberg, C., Chan, L., Bardin, T. & Singh, J. A. (2012) Developing a provisional definition of flare in patients with established gout. *Arthritis and Rheumatism*. 64(5), 1508–1517. doi: 10.1002/art.33483

Galea, S. & Tracy, M. (2007) Participation Rates in Epidemiologic Studies. *Annals of Epidemiology*. 17(9), 643–653. doi:10.1016/j.annepidem.2007.03.013

Garrow, A. P., Papageorgiou, A. C., Silman, A. J., Thomas, E., Jayson, M. I. & Macfarlane, G. J. (2000) Development and validation of a questionnaire to assess disabling foot pain. *Pain*. *85*(1-2), 107–13. doi:10.1016/S0304-3959(99)00263-8

Garrow, A. P., Silman, A. J. & Macfarlane, G. J. (2004) The Cheshire foot pain and disability survey: A population survey assessing prevalence and associations. *Pain*. 110(1-2), 378–384. doi:10.1016/j.pain.2004.04.019

Garson, D. G. (2013) *Hierarchial Linear Modeling Guide and Applications*. Thousand Oaks, California, SAGE Publications Inc.

Ghasemi, A. & Zahediasl, S. (2012) Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*. 10(2), 486–489. doi:10.5812/ijem.3505

Ghisletta, P., Renaud, O., Jacot, N. & Courvoisier, D. (2015) Linear Mixed-Effects and Latent Curve Models for Longitudinal Life Course Analyses. In: Burton-Jeangros, C., Cullati, S., & Sacker, A. (eds.) *Life Course Research and Social Policies*. Cham, Springer. 155-178.

Grace-Martin, K. (2019) Three Issues in Sample Size Estimates for Multilevel Models. Available from: https://www.theanalysisfactor.com/sample-size-multilevel-models/ Accessed: October 2019

Grant, B. F., Harford, T. C., Dawson, D. A., Chou, P. S. & Pickering, R. P. (1995) The alcohol use disorder and associated disabilities interview schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample. *Drug and Alcohol Dependence*. *39*(1), 37–44. doi:10.1016/0376-8716(95)01134-K

Green, C., Brazier, J. & Deverill, M. (2000) Valuing health-related quality of life. A review of health state valuation techniques. *PharmacoEconomics*. 17(2), 151–65. doi:10.2165/00019053-200017020-00004

Grenard, J. L., Munjas, B. A., Adams, J. L., Suttorp, M., Maglione, M., McGlynn, E. A. & Gellad, W. F. (2011) Depression and medication adherence in the treatment of chronic diseases in the United States: a meta-analysis. *Journal of General Internal Medicine*. 26(10), 1175–1182. doi:10.1007/s11606-011-1704-y

Grundy, E. & Holt, G. (2001) The socioeconomic status of older adults: How should we measure it in studies of health inequalities? *Journal of Epidemiology and Community Health*. 55(12), 895–904. doi:10.1136/jech.55.12.895

Gurka, M. J. (2006) Selecting the best linear mixed model under REML. *American Statistician*. *60*(1), 19–26. doi:10.1198/000313006X90396

Gustavson, K., Von Soest, T., Karevold, E. & Roysamb, E. (2012) Attrition and generalizability in longitudinal studies: Findings from a 15-year population-based study and a Monte Carlo simulation study. *BMC Public Health*. 12, 918. doi:10.1186/1471-2458-12-918

Hahn, E. A., Cella, D., Chassany, O., Fairclough, D. L., Wong, G. Y. & Hays, R. D. (2007) Precision of health-related quality-of-life data compared with other clinical measures. *Mayo Clinic proceedings*. 82(10), 1244–54. doi:10.4065/82.10.1244

Halpern, R., Fuldeore, M. J., Mody, R. R., Patel, P. A. & Mikuls, T. R. (2009) The effect of serum urate on gout flares and their associated costs: an administrative claims analysis. *Journal of clinical rheumatology*. 15(1), 3–7. doi: 10.1097/RHU.0b013e3181945d2c

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B.S., Hogdson, D.J. & Inger, R. (2018) A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*. 6, e4794. doi:10.7717/peerj.4794

Harrold, L. R., Mazor, K. M., Velten, S., Ockene, I. S. & Yood, R. A. (2010) Patients and providers view gout differently: A qualitative study. *Chronic Illness*. 6(4), 263–271. doi:10.1177/1742395310378761

Hassey, A., Gerret, D. & Wilson, A. (2001) A survey of validity and utility of electronic patient records in a general practice. *British Medical Journal*. *322*(7299), 1401–1405. doi: 10.1136/bmj.322.7299.1401

Hayes, V., Morris, J., Wolfe, C. & Morgan, M. (1995) The SF-36 health survey questionnaire: Is it suitable for use with older adults? *Age and Ageing*. 24 (2), 120–125. doi:10.1093/ageing/24.2.120

Hays, R. D. & Reeve, B. B. (2008) Measurement and Modeling of Health-Related Quality of Life. In: Heggenhougen, K. & Quah, S. (eds.) *International Encyclopaedia of Public Health*. San Diego, Academic Press. 241–252.

Haywood, K., Brett, J., Salek, S., Marlett, N., Penman, C., Shklarov, S., Norris, C., Santana, M. J. & Staniszewska, S. (2015) Patient and public engagement in health-related quality of life and patient-reported outcomes research: what is important and why should we care? Findings from the first ISOQOL patient engagement symposium. *Quality of Life Research*. 24(5), 1069–1076. doi: 10.1007/s11136-014-0796-3

Hediger, M.A., Johnson, R.J., Miyazaki, H. & Endou, H. (2005) Molecular Physiology of Urate Transport. *Physiology.* 20(2), 125–133. doi:10.1152/physiol.00039.2004

Heinze, G., Wallisch, C. & Dunkler, D. (2018) Variable selection – A review and recommendations for the practicing statistician. *Biometrical Journal*. 60(3), 431–449. doi:10.1002/bimj.201700067

Hemingway, H., Croft, P., Perel, P., Hayden, J. A., Abrams, K., Timmis, A., Briggs, A., Udumyan, R., Moons, K.G.M, Steyerberg, E.W., Roberts, I., Schroter, S. & Altman, D. (2013) Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes. *British Medical Journal*. 346, e5595. doi:10.1136/bmj.e5595

Hennessy, C. H., Moriarty, D. G., Zack, M. M., Scherr, P. A. & Brackbill, R. (1994) Measuring health-related quality of life for public health surveillance. *Public health reports*. 109(5), 665–672.

Hickey, A., Barker, M., McGee, H. & O'Boyle, C. (2005) Measuring health-related quality of life in older patient populations: A review of current approaches. *PharmacoEconomics*. 23(10), 971–993. doi:10.2165/00019053-200523100-00002

Higginson, I. & Carr, A. (2001) Using quality of life measures in the clinical setting. *British Medical Journal*. 322 (7297), 1297–1300. doi: 10.1136/bmj.322.7297.1297

Hilbe, J. M. (2007) *Negative Binomial Regression*. New York, Cambridge University Press.

Hirsch, J. D., Lee, S. J., Terkeltaub, R., Khanna, D., Singh, J., Sarkin, A., Harvey, J. & Kavanaugh, A. (2008) Evaluation of an instrument assessing influence of gout on health-related quality of life. *Journal of Rheumatology*. 35(12), 2406–2414. doi:10.3899/jrheum.080506

Hirsch, J. D., Terkeltaub, R., Khanna, D., Singh, J., Sarkin, A., Shieh, M., Kavanaugh, A. & Lee, S. J. (2010) Gout disease-specific quality of life and the association with gout characteristics. *Patient Related Outcome Measures*. Jul(1), 1–8. doi:10.2147/PROM.S8310

Hoenig, J. M. & Heisey, D. M. (2001) The Abuse of Power. *The American Statistician*. 55(1), 1-6. doi:10.1198/000313001300339897

Holla, J. F. M., van der Leeden, M., Heymans, M. W., Roorda, L. D., Bierma-Zeinstra, S. M. A., Boers, M., Lems, W. F., Steultjens, M. P. & Dekker, J. (2014) Three trajectories of activity limitations in early symptomatic knee osteoarthritis: a 5-year follow-up study. *Annals of the Rheumatic Diseases*. 73(7), 1369–1375. doi: 10.1136/annrheumdis-2012-202984

Hooper, D., Coughlan, J. & Mullen, M. R. (2008) Structural Equation Modelling: Guidelines for Determining Model Fit Structural Equation Modelling: Guidelines for Determining Model Fit. *The Electronic Journal of Business Research Methods*. *6*(1), 53–60.

Hox, J. & Stoel, R. D. (2005) Multilevel and SEM Approaches to Growth Curve Multilevel and SEM Approaches to Growth Curve Modeling. In: Everitt, B.S. & Howell, D.C (eds.). *Encyclopaedia of Statistics in Behavioural Science*. Chichester, John Wiley & Sons. 1296-1305.

Hu, L. & Bentler, P. M. (1998) Fit indices in covariance structure modeling: Sensitivity to underparameterized model specification. *Psychological Methods*. *3*(4), 424–453. doi:10.1037/1082-989X.3.4.424

Hu, L. T. & Bentler, P. M. (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 6(1), 1–55. doi:10.1080/10705519909540118

Hui, M., Carr, A., Cameron, S., Davenport, G., Doherty, M., Forrester, H., Jenkins, W., Jordan, K. M., Mallen, C. D., McDonald, T. M., Nuki, G., Pywell, A., Zhang, W. & Roddy, E. (2017) The British society for rheumatology guideline for the management of gout. *Rheumatology*. 56(7), e1–e20. doi: 10.1093/rheumatology/kex156

Hunt, I. M., Silman, A. J., Benjamin1, S., Mcbeth, J. & Macfarlane, G. J. (1999) The prevalence and associated features of chronic widespread pain in the community using the 'Manchester' definition of chronic widespread pain. *Rheumatology*. *38*(3), 275–279. doi:10.1093/rheumatology/38.3.275

Hunter, D. J., York, M., Chaisson, C. E., Woods, R., Niu, J. & Zhang, Y. (2006) Recent diuretic use and the risk of recurrent gout attacks: The online case-crossover gout study. *Journal of Rheumatology*. *33*(7), 1341–1345.

Hyndman, D., Liu, S. & Miner, J. N. (2016) Urate Handling in the Human Body. *Current Rheumatology Reports*. 18, 34. doi:10.1007/s11926-016-0587-7

Iannaccone, C.K., Fossel, A., Tsao, H., Cui, J.,Weinblatt, M. & Shadick, N. (2013) Factors associated with attrition in a longitudinal rheumatoid arthritis registry. *Arthritis Care & Research*. 65 (7), 1183–1189. doi:10.1002/acr.21940

Jackson, R., Shiozawa, A., Buysman, E. K., Altan, A., Korrer, S. & Choi, H. (2015) Flare frequency, healthcare resource utilisation and costs among patients with gout in a managed care setting: a retrospective medical claims-based analysis. *BMJ Open*. 5, e007214 1–9. doi: 10.1136/bmjopen-2014-007214

Jaeschke, R., Singer, J. & Guyatt, G. H. (1989) Ascertaining the Minimal Clinically Important Difference. *Controlled Clinical Trials*. 10, 407–415. doi:10.1016/0197-2456(89)90005-6

Jaffe, D. H., Eisenbach, Z., Neumark, Y. D. & Manor, O. (2005) Individual, household and neighborhood socioeconomic status and mortality: A study of absolute and relative deprivation. *Social Science and Medicine*. 60, 989–997. doi:10.1016/j.socscimed.2004.06.047

Janssen, C. A., Oude Voshaar, M. A. H., ten Klooster, P. M., Jansen, T. L. T. A., Vonkeman, H. E. & Van De Laar, M. A. F. J. (2019) A systematic literature review of patient-reported outcome measures used in gout: An evaluation of their content and measurement properties. *Health and Quality of Life Outcomes*. 17, 63, 1–11. doi:10.1186/s12955-019-1125-x

Janssens, H. J., Lucassen, P.L.B.J., Van der Laar, F., Janssen, M. & Van De Lisdonk, E. H. (2008) Systemic corticosteroids for acute gout (Review). *Cochrane Database of Systematic Reviews*. 2, CD005521. doi:10.1002/14651858.CD005521.pub2

Jeffries, N. O. (2003) A note on 'Testing the number of components in a normal mixture.' *Biometrika*. 90(4), 991–994. doi:10.1093/biomet/90.4.991

Jung, T. & Wickrama, K. A. (2008) An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*. 2(1), 302–317. doi: 10.1111/j.1751-9004.2007.00054.x

Juraschek, S. P., Miller, E.R. & Gelber, A. C. (2013) Body Mass Index, Obesity, and Prevalent Gout in the United States in 1988 – 1994 and 2007 – 2010. *Arthritis Care & Research*. 65(1), 127–132. doi:10.1002/acr.21791

Kapetanovic, M. C., Nilsson, P., Turesson, C., Englund, M., Dalbeth, N. & Jacobsson, L. (2018) The risk of clinically diagnosed gout by serum urate levels: Results from 30 years follow-up of the Malmö preventive project cohort in southern Sweden. *Arthritis Research & Therapy*. 20, 190. doi:10.1186/s13075-018-1697-6

Kempen, G. I. J. M., Jelicic, M. & Ormel, J. (1997) Personality, Chronic Medical Morbidity, and Health-Related Quality of Life among Older Persons. *Health Psychology*. 16(6), 539–546. doi:10.1037/0278-6133.16.6.539

Kenny, D. A. (2015). Measuring model fit. Available from: http://davidakenny.net/cm/fit.htm Accessed: February 2018

Keys, A., Fidanza, F., Karvonen, M.J., Kimura, N. & Taylor, H.L. (1972) Indices of relative weight and obesity. *Journal of Chronic Diseases*. 25 (6–7), 329-34.

Khanna, D., Ahmed, M., Yontz, D., Ginsburg, S.S., Park, G.S., Leonard, A. & Tsevat, J. (2008) The disutility of chronic gout. *Quality of Life Research*. 17 (5), 815–822. doi:10.1007/s11136-008-9355-0

Khanna, D., Fitzgerald, J. D., Khanna, P. P., Bae, S., Singh, M. K., Neogi, T., Pillinger, M. H., Merill, J., Lee, S., Prakash, S., Kaldas, M., Gogia, M., Perez-Ruiz, F., Taylor, W., Lioté, F., Choi, H., Singh, J. A., Dalbeth, N., Kaplan, S., Niyyar, V., Jones, D., Yarows, S. A., Roessler, B., Kerr, G., King, C., Levy, G., Furst, D. E., Edwards, N. L., Mandell, B., Schumacher, H. R., Robbins, M., Wenger, N. & Terkeltaub, R. (2012b) 2012 American college of rheumatology guidelines for management of gout. Part 1: Systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. *Arthritis Care & Research*. 64(10), 1431–1446. doi:10.1002/acr.21772

Khanna, D., Khanna, P. P., Fitzgerald, J. D., Singh, M. K., Bae, S., Neogi, T., Pillinger, M. H., Merill, J., Lee, S., Prakash, S., Kaldas, M., Gogia, M., Perez-Ruiz, F., Taylor, W., Lioté, F., Choi, H., Singh, J.A., Dalbeth, N., Kaplan, S., Niyyar, V., Jones, D., Yarows, S. A., Roessler, B., Kerr, G., King, C., Levy, G., Furst, D. E., Edwards, N. L., Mandell, B., Schumacher, H. R., Robbins, M., Wenger, N. & Terkeltaub, R. (2012a) 2012 American college of rheumatology guidelines for management of gout. Part 2: Therapy and antiinflammatory prophylaxis of acute gouty arthritis. *Arthritis Care & Research*. 64(10), 1447–1461. doi: 10.1002/acr.21773

Khanna, D., Sarkin, A. J., Khanna, P. P., Shieh, M. M., Kavanaugh, A. F., Terkeltaub, R. A., Lee, S. J., Singh, J. A. & Hirsch, J. D. (2011b) Minimally important differences of the gout impact scale in a randomized controlled trial. *Rheumatology*. 50(7), 1331–1336. doi: 10.1093/rheumatology/ker023

Khanna, P. P., Gladue, H. S., Singh, M. K., FitzGerald, J. D., Bae, S., Prakash, S., Kaldas, M., Gogia, M., Berrocal, V., Townsend, W., Terkeltaub, R. & Khanna, D. (2014) Treatment of acute gout: A systematic review. *Seminars in Arthritis and Rheumatism*. 44(1), 31–38. doi: 10.1016/j.semarthrit.2014.02.003

Khanna, P.P. & Khanna, D. (2012) Health-Related Quality of Life and Outcome Measures in Gout. In: Terkeltaub, R. (ed.) *Gout and Other Crystal Arthropathies*. Philadelphia, Elsevier. 217–225.

Khanna, P. P., Khanna, D., Shiozawa, A., Walker, V., Bancroft, T., Essoi, B. & Akhras, K. S. (2015) Health-related quality of life and treatment satisfaction in patients with gout: Results from a cross-sectional study in a managed care setting. *Patient Preference and Adherence*. 9, 971–981. doi:10.2147/PPA.S83700

Khanna, P.P., Nuki, G., Bardin, T., Tausche, A.-K., Forsythe, A., Goren, A., Vietri, J. T. & Khanna, D. (2012c) Tophi and frequent gout flares are associated with impairments to quality of life, productivity, and increased healthcare resource use: Results from a cross-sectional survey. *Health and Quality of Life Outcomes*. 10, 117. doi: 10.1186/1477-7525-10-117

Khanna, P. P., Perez-Ruiz, F., Maranian, P. & Khanna, D. (2011a) Long-term therapy for chronic gout results in clinically important improvements in the health-related quality of life: Short form-36 is responsive to change in chronic gout. *Rheumatology*. 50(4), 40–745. doi: 10.1093/rheumatology/keq346

Kirkwood, B.R. & Sterne, J. (2003) *Essential medical statistics*. 2nd edition. Oxford, Blackwell Science.

Kongsted, A., Kent, P., Axen, I., Downie, A. S. & Dunn, K. M. (2016) What have we learned from ten years of trajectory research in low back pain? *BMC Musculoskeletal Disorders*. 17, 220. doi:10.1186/s12891-016-1071-2

Krishnan, E., Sokka, T., Häkkinen, A., Hubert, H. & Hannonen, P. (2004) Normative Values for the Health Assessment Questionnaire Disability Index: Benchmarking Disability in the General Population. *Arthritis and Rheumatism*. 50(3), 953–960. doi:10.1002/art.20048

Kristman, V.L., Manno, M. & Côté, P. (2005) Methods to account for attrition in longitudinal data: Do they work? A simulation study. *European Journal of Epidemiology*. 20, 657–662. doi:10.1007/s10654-005-7919-7

Kroenke, K., Bair, M.J., Teresa M.S., Damush, M., Wu, J., Hoke, S., Sutherland, J. & Tu, W. (2009) Optimized Antidepressant Therapy and Pain Self-Management in Primary Care Patients with Depression and Musculoskeletal Pain: A Randomized Controlled Trial. *Journal of the American Medical Association.* 301(20), 2099–2110. doi:10.1001/jama.2009.723

Kroenke, K., Spitzer, R. L. & Williams, J. B. W. (2001) The PHQ-9 Validity of a Brief Depression Severity Measure. *Journal of General Internal Medicine*. *16*(9), 606–613. doi:10.1046/j.1525-1497.2001.016009606.x

Kuo, C.F., Grainge, M. J., Mallen, C., Zhang, W. & Doherty, M. (2015a) Rising burden of gout in the UK but continuing suboptimal management: A nationwide population study. *Annals of the Rheumatic Diseases*. 74(4), 661–667. doi:10.1136/annrheumdis-2013-204463

Kuo, C.F., Grainge, M. J., Mallen, C., Zhang, W. & Doherty, M. (2016) Comorbidities in patients with gout prior to and following diagnosis: case-control study. *Annals of the Rheumatic Diseases*. 75(1), 210–217. doi:10.1136/annrheumdis-2014-206410

Kuo, C.F., Grainge, M. J., See, L. C., Yu, K. H., Luo, S. F., Zhang, W. & Doherty, M. (2015c) Epidemiology and management of gout in Taiwan: A nationwide population study. *Arthritis Research & Therapy*. 17, 13. doi: 10.1186/s13075-015-0522-8

Kuo, C.F., Grainge, M. J., Zhang, W. & Doherty, M. (2015b) Global epidemiology of gout: Prevalence, incidence and risk factors. *Nature Reviews Rheumatology*. 11(11), 649–662. doi:10.1038/nrrheum.2015.91

Kydd, A., Seth, R., Buchbinder, R., Edwards, C. & Bombardier, C. (2014) Uricosuric medications for chronic gout (Review). *Cochrane Database of Systematic Reviews*. 11, CD010457. doi: 10.1002/14651858.CD010457.pub2

Lacey, R.J., Belcher, J., Rathod, T., Wilkie, R., Thomas, E. & McBeth, J. (2014) Pain at multiple body sites and health-related quality of life in older adults: Results from the north Staffordshire osteoarthritis project. *Rheumatology*. 53 (11), 2071–2079. doi:10.1093/rheumatology/keu240.

Lacey, R.J., Jordan, K.P. & Croft, P.R., (2013) Does attrition during follow-up of a population cohort study inevitably lead to biased estimates of health status? *PLoS ONE*. 8(12), e83948. doi: 10.1371/journal.pone.0083948

Lacey, R. J., Lewis, M., Jordan, K., Jinks, C., & Sim, J. (2005) Interrater reliability of scoring of pain drawings in a self-report health survey. *Spine*. 30(16), E455–E458. doi:10.1097/01.brs.0000174274.38485.ee

La-Crette, J., Jenkins, W., Fernandes, G., Valdes, A.M., Doherty, M. & Abhishek, A. (2018) First validation of the gout activity score against gout impact scale in a primary care based gout cohort. *Joint Bone Spine*. 85(3), 323–325. doi: 10.1016/j.jbspin.2017.05.010

Lambert, D. (1992) Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics*. *34*(1), 1–14. doi:10.1080/00401706.1992.10485228

Lee, A. H., Wang, K., Scott, J. A., Yau, K. K. W. & McLachlan, G. J. (2006) Multi-level zero-inflated poisson regression modelling of correlated count data with excess zeros. *Statistical Methods in Medical Research*. *15*(1), 47–61. doi: 10.1191/0962280206sm429oa

Lee, S. J., Hirsch, J. D., Terkeltaub, R., Khanna, D., Singh, J. A., Sarkin, A. & Kavanaugh, A. (2009) Perceptions of disease and health-related quality of life among patients with gout. *Rheumatology*. 48(5), 582–586. doi:10.1093/rheumatology/kep047

Lee, T. K., Wickrama, K. A. S. & O'Neal, C. W. (2017) Application of Latent Growth Curve Analysis With Categorical Responses in Social Behavioral Research. *Structural Equation Modeling: A Multidisciplinary Journal*. 25(2), 1-13. doi:10.1080/10705511.2017.1375858

Lee, W., Teng, G. G., Kok, J. C., Santosa, A., Lim, A. Y. N. & Wee, H. L. (2019) Validity and reliability of the Gout Impact Scale in a multi-ethnic Asian population. *International Journal of Rheumatic Diseases*. 22(8), 1427–1434. doi:10.1111/1756-185X.13595

Leffondré, K., Abrahamowicz, M., Regeasse, A., Hawker, G. A., Badley, E. M., McCusker, J. & Belzile, E. (2004) Statistical measures were proposed for identifying longitudinal patterns of

change in quantitative health indicators. *Journal of Clinical Epidemiology*. 57(10), 1049–1062. doi:10.1016/j.jclinepi.2004.02.012

Liddle, J., Roddy, E., Mallen, C. D., Hider, S. L., Prinjha, S., Ziebland, S. & Richardson, J. C. (2015) Mapping patients' experiences from initial symptoms to gout diagnosis: a qualitative exploration. *BMJ Open*. 5 (9), e008323. doi:10.1136/bmjopen-2015-008323

Lin, E. H. B., Katon, W., Von Korff, M., Tang, L., Williams, J. W., Kroenke, K., Hunkeler, E., Harpole, L., Hegel, M., Arean, P., Hoffing, M., Della Penna, R., Langston, C. & Unützer, J. (2003) Effect of Improving Depression Care on Pain and Functional Outcomes among Older Adults with Arthritis: A Randomized Controlled Trial. *Journal of the American Medical Association*. 290(18), 2428–2434. doi:10.1001/jama.290.18.2428

Lin, K.C., Lin, H.Y. & Chou, P. (2000) The interaction between uric acid level and other risk factors on the development of gout among asymptomatic hyperuricemic men in a prospective study. *Journal of Rheumatology*. 27(6), 1501–1505.

Lindsay, K., Gow, P., Vanderpyl, J., Logo, P. & Dalbeth, N. (2011) The experience and impact of living with gout: A study of men with chronic gout using a qualitative grounded theory approach. *Journal of Clinical Rheumatology*. 17(1), 1–6. doi: 10.1097/RHU.0b013e318204a8f9

Little, T. D., Jorgensen, T. D., Lang, K. M., & Moore, E. W. G. (2014) On the joys of missing data. *Journal of Pediatric Psychology*. 39(2), 151–162. doi:10.1093/jpepsy/jst048

Liu, H. & Powers, D. A. (2007) Growth curve models for zero-inflated count data: An application to smoking behavior. *Structural Equation Modeling*. *14*(2), 247–279. doi:10.1080/10705510709336746

Liu, X., Liu, K., Sun, Q., Wang, Y., Meng, J., Xu, Z. & Shi, Z. (2018) Efficacy and safety of febuxostat for treating hyperuricemia in patients with chronic kidney disease and in renal transplant recipients: A systematic review and meta-analysis. *Experimental and Therapeutic Medicine*. 16(3), 1859–1865. doi: 10.3892/etm.2018.6367

Lo, Y., Mendell, N. R. & Rubin, D. B. (2001) Testing the Number of Components in a Normal Mixture. *Biometrika*. *88*(3), 767–778. doi:10.1093/biomet/88.3.767

Locascio, J. J. & Atri, A. (2011) An Overview of Longitudinal Data Analysis Methods for Neurological Research. *Dementia and Geriatric Cognitive Disorders Extra*. *1*(1), 330–357. doi:10.1159/000330228

Long, J. (2012) *Longitudinal Analysis for the Behavioural Sciences Using R*. Thousand Oaks, California, SAGE Publications Inc.

López López, C. O., Lugo, E. F., Alvarez-Hernández, E., Peláez-Ballestas, I., Burgos-Vargas, R. & Vázquez-Mellado, J. (2017) Severe tophaceous gout and disability: changes in the past 15 years. *Clinical Rheumatology*. 36(1), 199–204. doi:10.1007/s10067-016-3381-z

Lyons, R. A., Wareham, K., Lucas, M., Price, D., Williams, J. & Hutchings, H. A. (1999) SF-36 scores vary by method of administration: Implications for study design. *Journal of Public Health Medicine*. *21*(1), 41–45. doi:10.1093/pubmed/21.1.41

Maas, C. J. M. & Hox, J. J. (2004) Robustness issues in multilevel regression analysis. *Statistica Neerlandica*. 58 (2), 127–137. doi:10.1046/j.0039-0402.2003.00252.x

MacCallum, R. C., Browne, M. W. & Sugawara, H. M. (1996) Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*. *1*(2), 130–149. doi:10.1037/1082-989X.1.2.130

MacFarlane, L.A., Liu, C., Solomon, D.C. & Kim, S. C. (2016) Validation of claims-based algorithms for gout flares. *Pharmacoepidemiology and Drug Safety*. *25 (7)*, 820–826. doi: 10.1002/pds.4044

MacKenzie, C. R. (2015) Gout and Hyperuricemia: an Historical Perspective. *Current Treatment Options in Rheumatology*. 1, 119–130. doi:10.1007/s40674-015-0012-9

Major, T. J., Dalbeth, N., Stahl, E. A. & Merriman, T. R. (2018) An update on the genetics of hyperuricaemia and gout. *Nature Reviews Rheumatology*. 14(6), 351–353. doi: 10.1038/s41584-018-0004-x

Man, C. Y., Cheung, I. T. F., Cameron, P. A. & Rainer, T. H. (2007) Comparison of Oral Prednisolone/Paracetamol and Oral Indomethacin/Paracetamol Combination Therapy in the Treatment of Acute Goutlike Arthritis: A Double-Blind, Randomized, Controlled Trial. *Annals of Emergency Medicine*. 49(5), 670–677. doi: 10.1016/j.annemergmed.2006.11.014

Mandal, A. K. & Mount, D. B. (2015) The molecular physiology of uric acid homeostasis. *Annual review of physiology*. 77, 323–45. doi: 10.1146/annurev-physiol-021113-170343

Manini, T. (2012) Development of Physical Disability in Older Adults. *Current Aging Science*. 4(3), 184–191. doi:10.2174/1874609811104030184

Martinon, F., Pétrilli, V., Mayor, A., Tardivel, A. & Tschopp, J. (2006) Gout-associated uric acid crystals activate the NALP3 inflammasome. *Nature*. 440(7081), 237–241. doi: 10.1038/nature04516

Mathes, T., Jaschinski, T., & Pieper, D. (2014) Adherence influencing factors - a systematic review of systematic reviews. *Archives of Public Health*. 72, 37. doi:10.1186/2049-3258-72-37

Matthews, F.E., Chatfield, M., Freeman, C., McCracken, C. & Brayne, C. (2004) Attrition and bias in the MRC cognitive function and ageing study: an epidemiological investigation. *BMC Public Health*. 4, 12. doi:10.1186/1471-2458-4-12

McAdams, M. A, Van Dam, R. M. & Hu, F. B. (2007) Comparison of self-reported and measured BMI as correlates of disease markers in US adults. *Obesity. 15*(1), 188–196. doi:10.1038/oby.2007.504

McColl, E. (2004) Best practice in symptom assessment: A review. *Gut*. 53(s4), 49–54. doi:10.1136/gut.2003.034355

McHorney, C. A, Ware, J. E. & Raczek, A. E. (1993) The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care*. *31*(3), 247–263. doi:10.2307/3765819

McLachlan, G.J. & Peel, D. (2000) *Finite Mixture Models*. New York, John Wiley & Sons.

Meier, C. R. & Jick, H. (1997) Omeprazole, other antiulcer drugs and newly diagnosed gout. *British Journal of Clinical Pharmacology*. 44(2), 175–178. doi:10.1046/j.1365-2125.1997.00647.x

Merriman, T. R., Choi, H. K. & Dalbeth, N. (2014) The genetic basis of gout. *Rheumatic Disease Clinics of North America*. 40(2), 279–290. doi:10.1016/j.rdc.2014.01.009

Middel, B. & Van Sonderen, E. (2002) Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. *International Journal of Integrated Care*. 2. doi:10.5334/ijic.65/

Mikuls, T. R., Farrar, J. T., Bilker, W. B., Fernandes, S., Schumacher, H. R. & Saag, K. G. (2005) Gout epidemiology: Results from the UK General Practice Research Database, 1990-1999. *Annals of the Rheumatic Diseases*. 64(2), 267–272. doi: 10.1136/ard.2004.024091

Mitnala, S., Phipps-Green, A., Franklin, C., Horne, A., Stamp, L. K., Merriman, T. R. & Dalbeth, N. (2016) Clinical and genetic features of diuretic-associated gout: A case-control study. *Rheumatology*. 55(7), 1172–1176. doi: 10.1093/rheumatology/kew018

Moi, J., Sriranganathan, M., Edwards, C. & Buchbinder, R. (2013) Lifestyle interventions for chronic gout. *Cochrane Database of Systematic Reviews*. 5, CD010039. doi: 10.1002/14651858.CD010039.pub2

Molto, L., AnnaGossec, S., Tezenas du Montcel, S., Vanier, A., Wendling, D. & Dougados, M. (2017) Disease activity trajectories in early axial spondyloarthritis: results from the DESIR cohort. *Annals of the Rheumatic Diseases*. 76(6), 1036–1041. doi.org/doi: 10.1136/annrheumdis-2016-209785

Moons, P. (2004) Why call it health-related quality of life when you mean perceived health status? *European Journal of Cardiovascular Nursing*. 3(4), 275–277. doi: 10.1016/j.ejcnurse.2004.09.004

Moons, P., Budts, W. & De Geest, S. (2006) Critique on the conceptualisation of quality of life: A review and evaluation of different conceptual approaches. *International Journal of Nursing Studies*. 43(7), 891–901. doi: 10.1016/j.ijnurstu.2006.03.015

Moore, A., Grime, J., Campbell, P. & Richardson, J. (2013) Troubling stoicism: Sociocultural influences and applications to health and illness behaviour. *Health*. 17(2), 159–173. doi:10.1177/1363459312451179

Moriarty, A. S., Gilbody, S., McMillan, D. & Manea, L. (2015) Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): A meta-analysis. *General Hospital Psychiatry*. *37*(6), 567–576. doi:10.1016/j.genhosppsych.2015.06.012

Morris, S. B. (2000) Distribution of the standardized mean change effect size for meta-analysis on repeated measures. *British Journal of Mathematical and Statistical Psychology*. 53, 17–29. doi:10.1348/000711000159150

Morris, S. B. & DeShon, R. P. (2002) Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*. 7(1), 105–125. doi:10.1037/1082-989X.7.1.105

Morton, S. M. B., Bandara, D. K., Robinson, E. M., & Atatoa Carr, P. E. (2012) In the 21st Century, what is an acceptable response rate? *Australian and New Zealand Journal of Public Health*. 36(2), 106–108. doi:10.1111/j.1753-6405.2012.00854.x

Moss-Morris, R., Weinman, J., Petrie, K. J., Horne, R., Cameron, L. D. & Buick, D. (2002) The Revised Illness Perception Questionnaire (IPQ-R). *Psychology and Health*. *17*(1), 1–16. doi:10.1080/08870440290001494

Muldoon, M. F., Barger, S. D., Flory, J. D. & Manuck, S. B. (1998) What are quality of life measurements measuring? *British Medical Journal*. 316(7130), 542–5. doi:10.1136/bmj.316.7130.542

Muth, C., Bales, K. L., Hinde, K., Maninger, N., Mendoza, S. P. & Ferrer, E. (2016) Alternative Models for Small Samples in Psychological Research: Applying Linear Mixed Effects Models and Generalized Estimating Equations to Repeated Measures Data. *Educational and Psychological Measurement*. *76*(1), 64–87. doi:10.1177/0013164415580432

Muthen, B. O. (2001) Second-generation structural equation modeling with a combination of categorical and continuous latent variables. New opportunities for latent class-latent growth modeling. In: Collins, L.M. & Sayers, A.G. (eds.). *New methods for the analysis of change*. Washington DC, American Psychological Association. 291–322

Muthén, B. O. & Muthén, L. K. (2000) Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*. *24*(6), 882–891. doi:10.1111/j.1530-0277.2000.tb02070.x

Muthén, L. K. & Muthén, B. O. (2002) How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power. *Structural Equation Modeling: A Multidisciplinary Journal*. *9*(4), 599–620. doi:10.1207/S15328007SEM0904

Muthén, L. K. & Muthén, B. O. (2017) *Mplus Statistical Analysis with Latent Variables User's Guide*. Version 8. Los Angeles CA, Muthen & Muthen.

Myck, M., Najsztub, M. & Oczkowska, M. (2019) Implications of Social and Material Deprivation for Changes in Health of Older People. *Journal of Aging and Health*. 1–25. doi:10.1177/0898264319826417

Nagin, D.S. (1999) Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods*. 4(2), 139–157. doi:10.1037/1082-989X.4.2.139

Nagin, D.S. & Land, K. C. (1993) Age, Criminal Careers, and Population Heterogeneity: Specification and Estimation of a Nonparametric, Mixed Poisson Model. *Criminology*. *31*(3), 327–362. doi:10.1111/j.1745-9125.1993.tb01133.x

Nagin, D.S. & Odgers, C. L. (2010) Group-Based Trajectory Modeling in Clinical Research. *Annual Review of Clinical Psychology. 6*, 109–138. doi:10.1146/annurev.clinpsy.121208.131413

National Health Service. (2019) The NHS Long Term Plan. Available at : https://doi.org/10.12968/jprp.2019.1.3.114 Accessed: September 2019

National Institute for Health and Care Excellence. (2008) Febuxostat for the management of hyperuricaemia in people with gout. (NICE Technology Appraisal Guidance 164). Available from: https://www.nice.org.uk/guidance/ta164 Accessed: November 2019

Neogi, T., Hunter, D. J., Chaisson, C. E., Allensworth-Davies, D. & Zhang, Y. (2006) Frequency and predictors of inappropriate management of recurrent gout attacks in a longitudinal study. *Journal of Rheumatology*. *33*(1), 104–109.

Nguyen, U. D. T., Zhang, Y., Louie-gao, Q., Niu, J., Felson, D. T., Lavalley, M. P. & Choi, H. K. (2017) The Obesity Paradox in Recurrent Attacks of Gout in Observational Studies: Clarification and Remedy. *Arthritis Care & Research*. 69(4), 561–566. doi: 10.1002/acr.22954

Nicholls, E., Thomas, E., van der Windt, D. A., Croft, P. R. & Peat, G. (2014) Pain trajectory groups in persons with, or at high risk of, knee osteoarthritis: Findings from the Knee clinical assessment study and the osteoarthritis initiative. *Osteoarthritis and Cartilage*. 22(12), 2041–2050. doi:10.1016/j.joca.2014.09.026

Nielsen, S. M., Bartels, E. M., Henriksen, M., Wæhrens, E. E., Gudbergsen, H., Bliddal, H., Astrup, A., Knop, F. K., Carmona, L., Taylor, W. J., Singh, J. A., Perez-Ruiz, F., Kristensen, L. E. & Christensen, R. (2017) Weight loss for overweight and obese individuals with gout: A systematic review of longitudinal studies. *Annals of the Rheumatic Diseases*. 76(11), 1870–1882. doi: 10.1136/annrheumdis-2017-211472

Nuki, G. & Simkin, P. A. (2006) A concise history of gout and hyperuricemia and their treatment. *Arthritis Research & Therapy*. 8, S1. doi:10.1186/ar1906

Nylund, K. L., Asparouhov, T. & Muthén, B. O. (2007) Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*. *14*(4), 535–569. doi:10.1080/10705510701575396

Obidoa, C., Reisine, S. & Cherniack, M. (2010) How Does the SF-36 Perform in Healthy Populations? A Structured Review of Longitudinal Studies. *Journal of Social, Behavioral, and Health Sciences*. 4(1), 30–48. doi: 10.5590/JSBHS.2010.04.1.02

Office for National Statistics. (2018) *Living longer - how our population is changing and why it matters*. Office for National Statistics, London.

Ogollah, R. O., Konstantinou, K., Stynes, S. & Dunn, K. M. (2018) Determining One-Year Trajectories of Low-Back–Related Leg Pain in Primary Care Patients: Growth Mixture Modeling of a Prospective Cohort Study. *Arthritis Care & Research.* 70(12), 1840–1848. doi: 10.1002/acr.23556

Osborne, J. (2013) *Best Practices in Data Cleaning*. London, SAGE Publications.

Pagels, A.A., Soferkvist, B.K., Medin, C., Hylander, B. & Heiwe, S. (2012) Health-related quality of life in different stages of chronic kidney disease and at initiation of dialysis treatment. *Health and Quality of Life Outcomes*. 10, 71. doi: 10.1186/1477-7525-10-71

Pallant, J. (2016) *SPSS Survival Manual*. 6th edition. Maidenhead, Open University Press.

Paz, S. H., Liu, H., Fongwa, M. N., Morales, L. S. & Hays, R. D. (2009) Readability estimates for commonly used health-related quality of life surveys. *Quality of Life Research*. 18, 889–900. doi:10.1007/s11136-009-9506-y

Peters, S. A. E., Bots, M. L., Den Ruijter, H. M., Palmer, M. K., Grobbee, D. E., Crouse, J. R., O'Leary, D.H., Evans, G.W., Raichlen, J.S., Moons, K.G. & Koffijberg, H. (2012) Multiple

imputation of missing repeated outcome measurements did not add to linear mixed-effects models. *Journal of Clinical Epidemiology*. 65(6), 686–695. doi: 10.1016/j.jclinepi.2011.11.012

Petrie, A. & Sabin, C. (2009) *Medical Statistics at a Glance*. 3rd edition. Oxford, John Wiley & Sons.

Pétrilli, V. & Martinon, F. (2007) The inflammasome, autoinflammatory diseases, and gout. *Joint Bone Spine*. 74(6), 571–576. doi:10.1016/j.jbspin.2007.04.004

Peugh, J. & Fan, X. (2012) How Well Does Growth Mixture Modeling Identify Heterogeneous Growth Trajectories? A Simulation Study Examining GMM's Performance Characteristics. *Structural Equation Modeling*. *19*(2), 204–226. doi:10.1080/10705511.2012.659618

Peyre, H., Leplège, A. & Coste, J. (2011) Missing data methods for dealing with missing items in quality of life questionnaires. A comparison by simulation of personal mean score, full information maximum likelihood, multiple imputation, and hot deck techniques applied to the SF-36 in the French 2003 decennial health survey. *Quality of Life Research*. 20, 287–300. doi:10.1007/s11136-010-9740-3

Porcheret, M., Hughes, R., Evans, D., Jordan, K., Whitehurst, T., Ogden, H. & Croft, P. (2004) Data Quality of General Practice Electronic Health Records: The Impact of a Program of Assessments, Feedback, and Training. *Journal of Medical Information Association*. 11(1), 78–86. doi:10.1197/jamia.M1362.As

Porter, R. & Rousseau, G. S. (1998) *Gout The Patrician Malady*. New Haven, Yale University Press.

Prior, J. A., Mallen, C. D., Chandratre, P., Muller, S., Richardson, J. & Roddy, E. (2016) Gout characteristics associate with depression, but not anxiety, in primary care: Baseline findings from a prospective cohort study. *Joint Bone Spine*. *83*(5), 553–558. doi:10.1016/j.jbspin.2015.10.008

Protheroe, J., Whittle, R., Bartlam, B., Estacio, E. V., Clark, L. & Kurth, J. (2017) Health literacy, associated lifestyle and demographic factors in adult population of an English city: a cross-sectional survey. *Health Expectations*. 20(1), 112–119. doi:10.1111/hex.12440

Proudman, C., Lester, S.E., Gonzalez-Chica, D.A., Gill, T.K., Dalbeth, N. & Hill, C.L. (2019) Gout, flares, and allopurinol use: A population-based study. *Arthritis Research & Therapy*. 21, 132. doi:10.1186/s13075-019-1918-7

Prowse, R. L., Dalbeth, N., Kavanaugh, A., Adebajo, A. O., Gaffo, A. L., Terkeltaub, R., Mandell, B. F., Suryana, B. P. P., Goldenstein-Schainberg, C., Diaz-Torne, C., Khanna, D., Lioté, F., McCarthy, G., Kerr, G. S., Yamanaka, H., Janssens, H., Baraf, H. F., Chen, J. H., Vazquez-Mellado, J., Harrold, L. R., Stamp, L. K., Van De Laar, M. A., Janssen, M., Doherty, M., Boers, M., Edwards, N. L., Gow, P., Chapman, P., Khanna, P., Helliwell, P. S., Grainger, R., Schumacher, H. R., Neogi, T., Jansen, T. L., Louthrenoo, W., Sivera, F. & Taylor, W. J. (2013) A Delphi exercise to identify characteristic features of gout - Opinions from patients and physicians, the first stage in developing new classification criteria. *Journal of Rheumatology*. 40(4), 498–505. doi: 10.3899/jrheum.121037

Public Health England. (2015) *Improving health literacy to reduce health inequalities.* Available from: https://www.gov.uk/government/publications/local-action-on-health-inequalities-improving-health-literacy Accessed: September 2019

Rabe-Hesketh, S. & Skrondal, A. (2008) *Multilevel and Longitudinal Modeling Using Stata*. 2nd edition. College Station, Stata Press.

Rai, S. K., Aviña-Zubieta, J. A., McCormick, N., De Vera, M. A., Shojania, K., Sayre, E. C. & Choi, H. K. (2017a) The rising prevalence and incidence of gout in British Columbia, Canada: Population-based trends from 2000 to 2012. *Seminars in Arthritis and Rheumatism*. 46(4), 451–456. doi:10.1016/j.semarthrit.2016.08.006

Rai, S. K., Fung, T. T., Lu, N., Keller, S. F., Curhan, G. C. & Choi, H. K. (2017b) The Dietary Approaches to Stop Hypertension (DASH) diet, Western diet, and risk of gout in men: Prospective cohort study. *British Medical Journal*. 357, j1794. doi: 10.1136/bmj.j1794

Ram, N. & Grimm, K. J. (2009) Methods and Measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development*. *33*(6), 565–576. doi:10.1177/0165025409343765

Rashid, N., Levy, G., Wu, Y.L., Zheng, C., Koblick, R. & Cheetham, T. (2015) Patient and clinical characteristics associated with gout flares in an integrated healthcare system. *Rheumatology International*. 35(11), 1799–1807. doi:10.1007/s00296-015-3284-3

Rees, F., Jenkins, W. & Doherty, M. (2013) Patients with gout adhere to curative treatment if informed appropriately: Proof-of-concept observational study. *Annals of the Rheumatic Diseases*. 72(6), 826–830. doi: 10.1136/annrheumdis-2012-201676

Reinecke, J. & Seddig, D. (2011) Growth mixture models in longitudinal research. *Advances in Statistical Analysis*. 95 (4), 415–434. doi:10.1007/s10182-011-0171-4

Richardson, J. C., Liddle, J., Mallen, C. D., Roddy, E., Prinjha, S., Ziebland, S. & Hider, S. (2015) "Why me? I don't fit the mould ... I am a freak of nature": A qualitative study of women's experience of gout. *BMC Women's Health.* 15, 122. doi:10.1186/s12905-015-0277-z

Richette, P., Doherty, M., Pascual, E., Barskova, V., Becce, F., Castañeda-Sanabria, J., Coyfish, M., Guillo, S., Jansen, T. L., Janssens, H., Lioté, F., Mallen, C.D., Nuki, G., Perez-Ruiz, F., Pimentao, J., Punzi, L., Pywell, T., So, A., Tausche, A. K., Uhlig, T., Zavada, J., Zhang, W., Tubach, F. & Bardin, T. (2017) 2016 updated EULAR evidence-based recommendations for the management of gout. *Annals of the Rheumatic Diseases*. 76(1), 29–42. doi:10.1136/annrheumdis-2016-209707

Richette, P., Doherty, M., Pascual, E., Barskova, V., Becce, F., Castañeda-Sanabria, J., Coyfish, M., Guillo, S., Jansen, T., Janssens, H., Lioté, F., Mallen, C.D., Nuki, G., Perez-Ruiz, F., Pimentao, J., Punzi, L., Pywell, A., So, A., Tausche, A.K., Uhlig, T., Zavada, J., Zhang, W., Tubach, F. & Bardin, T. (2020) 2018 updated European League Against Rheumatism evidence-based recommendations for the diagnosis of gout. *Annals of the Rheumatic Diseases*. 79(1), 31-38. doi:10.1136/annrheumdis-2019-215315

Riley, R., Hayden, J.A., Steyerberg, E.W., Moons, K.G.M., Abrams, K., Kyzas, P.A., Malats, N., Briggs, A., Schroter, S., Altman, D.G. & Hemingway, H. (2013) Prognosis Research Strategy (PROGRESS) 2: Prognostic Factor Research. *PLOS Medicine*. 10(2): e1001380. doi:10.1371/journal.pmed.1001380

Robinson, P. C., Taylor, W. J. & Dalbeth, N. (2015) An observational study of gout prevalence and quality of care in a national Australian general practice population. *Journal of Rheumatology*. 42(9), 1702–1707. doi: 10.3899/jrheum.150310

Roch-Ramel, F. & Guisan, B. (1999) Renal Transport of Urate in Humans. *News in physiological sciences*. 14(April), 80–84.

Roddy, E. (2011) Revisiting the pathogenesis of podagra: why does gout target the foot? *Journal of Foot and Ankle Research*. 4, 13. doi: 10.1186/1757-1146-4-13

Roddy, E. & Choi, H. (2014) Epidemiology of Gout. *Rheumatic Disease Clinics of North America*. 40(2), 155–175. doi: 10.1016/j.rdc.2014.01.001

Roddy, E., Clarkson, K., Blagojevic-Bucknall, M., Mehta, R., Oppong , R., Avery , A., Hay , E.M., Heneghan, C., Hartshorne, L., Hooper, J., Hughes, G., Jowett, S., Lewis, M., Little, P., McCarterney, K., Mahatani, K.R., Nunan, D., Santer, M., Williams, S. & Mallen, C.D. (2019) Open-label randomised pragmatic trial (CONTACT) comparing naproxen and low-dose colchicine for the treatment of gout flares in primary care. *Annals of the Rheumatic Diseases*. doi:10.1136/annrheumdis-2019-216154.

Roddy, E., Mallen, C. D. & Doherty, M. (2013) Gout. *British Medical Journal*. *347*, f5648. doi.org/10.1136/bmj.f5648

Roddy, E., Mallen, C. D., Hider, S. L. & Jordan, K. P. (2010) Prescription and comorbidity screening following consultation for acute gout in primary care. *Rheumatology*. 49(1), 105–111. doi:10.1093/rheumatology/kep332

Roddy, E., Menon, A., Hall, A., Datta, P. & Packham, J. (2013) Polyarticular sonographic assessment of gout: a hospital-based cross-sectional study. *Joint Bone Spine*. 80(3), 295–300. doi: 10.1016/j.jbspin.2012.09.017

Roddy, E., Muller, S., Rome, K., Chandratre, P., Hider, S. L., Richardson, J., Blagojevic-Bucknall, M. & Mallen, C. D. (2015) Foot problems in people with gout in primary care: baseline findings from a prospective cohort study. *Journal of Foot and Ankle Research*. *8*(1), 31. doi:10.1186/s13047-015-0090-9

Roddy, E., Zhang, W. & Doherty, M. (2007a) Are joints affected by gout also affected by osteoarthritis? *Annals of Rheumatic Diseases*. 66(10), 1374–7. doi:10.1136/ard.2006.063768

Roddy, E., Zhang, W. & Doherty, M. (2007b) Concordance of the management of chronic gout in a UK primary-care population with the EULAR gout recommendations. *Annals of Rheumatic Diseases*. 66(10), 1311–1315. doi:10.1136/ard.2007.070755

Roddy, E., Zhang, W. & Doherty, M. (2007c) Is gout associated with reduced quality of life? A case-control study. *Rheumatology*. *46*(9), 1441–1444. doi:10.1093/rheumatology/kem150

Roddy, E., Zhang, W. & Doherty, M. (2007d) Validation of a self-report instrument for assessment of hallux valgus. *Osteoarthritis and Cartilage*. *15*(9), 1008–1012. doi:10.1016/j.joca.2007.02.016

Rodríguez, L.A.G., Cea Soriano, L. & Choi, H.K. (2010) Impact of Diabetes Against the Future Risk of Developing Gout. *Annals of the Rheumatic Diseases*. 69(12), 2090–2094. doi: 10.1136/ard.2010.130013

Rothenbacher, D., Primatesta, P., Ferreira, A., Cea-Soriano, L. & Rodríguez, L. A. G. (2011) Frequency and risk factors of gout flares in a large population-based cohort of incident gout. *Rheumatology*. 50(5), 973–981. doi:10.1093/rheumatology/keq363

Roughley, M. J., Belcher, J., Mallen, C. D. & Roddy, E. (2015) Gout and risk of chronic kidney disease and nephrolithiasis: Meta-analysis of observational studies. *Arthritis Research & Therapy*. 17, 90. doi:10.1186/s13075-015-0610-9PMCID: PMC4404569

Roughley, M., Sultan, A. A., Clarson, L., Muller, S., Whittle, R., Belcher, J., Mallen, C. D. & Roddy, E. (2018) Risk of chronic kidney disease in patients with gout and the impact of urate lowering therapy: a population-based cohort study. *Arthritis Research & Therapy*. 20, 243. doi:10.1186/s13075-015-0610-9

Royston, P., Moons, K. G. M., Altman, D. G. & Vergouwe, Y. (2009) Prognosis and prognostic research: Developing a prognostic model. *British Medical Journal. 338*, b604. doi:10.1136/bmj.b604

Rzewuska, M., Mallen, C. D., Strauss, V. Y., Belcher, J. & Peat, G. (2015) One-year trajectories of depression and anxiety symptoms in older patients presenting in general practice with musculoskeletal pain: A latent class growth analysis. *Journal of Psychosomatic Research*. 79(3), 195–201. doi:10.1016/j.jpsychores.2015.05.016

Sahlqvist, S., Song, Y., Bull, F., Adams, E., Preston, J. & Ogilvie, D. (2011) Effect of questionnaire length, personalisation and reminder type on response rate to a complex postal survey: randomised controlled trial. *BMC Medical Research Methodology*. 11, 62. doi: 10.1186/1471-2288-11-62

Sangha, O., Stucki, G., Liang, M. H., Fossel, A. H. & Katz, J. N. (2003) The Self-Administered Comorbidity Questionnaire: a new method to assess comorbidity for clinical and health services research. *Arthritis and Rheumatism*. *49*(2), 156–163. doi:10.1002/art.10993

Sarawate, C. A., Patel, P. A., Schumacher, H., Yang, W., Brewer, K. K. & Bakst, A. W. (2006) Serum Urate Levels and Gout Flares. *Journal of Clinical Pharmacy and Therapeutics*. 12(2), 61–65. doi:10.1097/01.rhu.0000209882.50228.9f

Sarkin, A. J., Levack, A. E., Shieh, M. M., Kavanaugh, A. F., Khanna, D., Singh, J. A., Terkeltaub, R. A., Lee, S. J. & Hirsch, J. D. (2010) Predictors of doctor-rated and patient-rated gout severity: Gout impact scales improve assessment. *Journal of Evaluation in Clinical Practice*. 16(6), 1244–1247. doi:10.1111/j.1365-2753.2009.01303.x

Saseen, J.J., Agashivala, N., Allen, R.R., Ghushchyan, V., Yadao, A.M. & Nair, K.V. (2012) Comparison of patient characteristics and gout-related health-care resource utilization and costs in patients with frequent versus infrequent gouty arthritis attacks. *Rheumatology*. 51 (11), 2004–2012. doi:10.1093/rheumatology/kes183

Schafer, JL. (1999) Multiple imputation: a primer. *Statistical Methods in Medical Research*. 8 (1), 3–15. doi:10.1177/096228029900800102

Scheepers, L. E. J. M., Burden, A. M., Arts, I. C. W., Spaetgens, B., Souverein, P., De Vries, F. & Boonen, A. (2018) Medication adherence among gout patients initiated allopurinol: A retrospective cohort study in the Clinical Practice Research Datalink (CPRD). *Rheumatology*. 57(9), 1641–1650. doi: 10.1093/rheumatology/key155

Schmidt, C.O., Raspe, H., Pfingsten, M., Hasenbring, M., Dieter Basler, H., Eich,W. & Kohlmann, T. (2011) Does attrition bias longitudinal population-based studies on back pain? *European Journal of Pain*. 15 (1), 84–91. doi:10.1016/j.ejpain.2010.05.007

Schumacher, H. R., Taylor, W., Edwards, L., Grainger, R., Schlesinger, N., Dalbeth, N., Sivera, F., Singh, J., Evans, R., Waltrip, R. W., Diaz-Torne, C., Macdonald, P., Mcqueen, F. & Perez-Ruiz, F. (2009) Outcome domains for studies of acute and chronic gout. *Journal of Rheumatology*. 36(10), 2342–2345. doi:10.3899/jrheum.090370

Scire, C. A, Manara, M., Cimmino, M. a, Govoni, M., Salaffi, F., Punzi, L., Monti, M.C., Carra, G., Montecucco, C., Matucci-Cerinic, M. & Minisola, G. (2013) Gout impacts on function and health-related quality of life beyond associated risk factors and medical conditions: results from the KING observational study of the Italian Society for Rheumatology (SIR). *Arthritis research & therapy*. 15, R101. doi:10.1186/ar4281

Sedgwick, P. (2014) Pitfalls of statistical hypothesis testing: Multiple testing. *British Medical Journal*. 349, 5310. doi:10.1136/bmj.g5310

Seidel, J. A., Miller, S. D. & Chow, D. L. (2014) Effect size calculations for the clinician: Methods and comparability. *Psychotherapy Research*. 24 (4), 470-484. doi:10.1080/10503307.2013.840812

Seth, R., Kydd, A. S., Buchbinder, R., Bombardier, C. & Edwards, C. J. (2014) Allopurinol for chronic gout review. *Cochrane Database of Systematic Reviews*. 10, CD006077. doi:10.1002/14651858.CD006077.pub3

Shih, T.H. & Fan, X. (2008) Comparing Response Rates from Web and Mail Surveys: A Meta-Analysis. *Field Methods*. *20*(3), 249–271. doi:10.1177/1525822X08317085

Shippee, N. D., Domecq Garces, J. P., Prutsky Lopez, G. J., Wang, Z., Elraiyah, T. A., Nabhan, M., Brito, J.P., Boehmer, K., Hasan, R., Firwana, B., Erwin, P. J., Montori, V. M. & Murad, M. H. (2015) Patient and service user engagement in research: A systematic review and synthesized framework. *Health Expectations*. *18*(5), 1151–1166. doi:10.1111/hex.12090

Shiyko, M. P., Li, Y. & Rindskopf, D. (2012) Poisson growth mixture modeling of intensive longitudinal data: An application to smoking cessation behavior. *Structural Equation Modeling*. *19*(1), 65–85. doi:10.1080/10705511.2012.634722

Siemons, L., ten Klooster, P. M., Vonkeman, H. E., Glas, C. A. W. & Van De Laar, M. A. F. J. (2014) Distinct trajectories of disease activity over the first year in early rheumatoid arthritis patients following a treat-to-target strategy. *Arthritis Care & Research*. 66(4), 625–630. doi:10.1002/acr.22175

Sim, J. & Wright, C. (2000) *Research in Health Care Concepts, Designs and Methods*. Cheltenham, Nelson Thornes Ltd.

Simpson, C. F., Boyd, C. M., Carlson, M. C., Griswold, M. E., Guralnik, J. M. & Fried, L. P. (2004) Agreement between Self-Report of Disease Diagnoses and Medical Record Validation in Disabled Older Women: Factors That Modify Agreement. *Journal of the American Geriatrics Society*. *52*(1), 123–127. doi:10.1111/j.1532-5415.2004.52021.x

Sinclair, M., O'Toole, J., Malawaraarachchi, M. & Leder, K. (2012) Comparison of response rates and cost-effectiveness for a community-based survey: postal, internet and telephone modes with generic or personalised recruitment approaches. *BMC Medical Research Methodology*. *12*, 132. doi:10.1186/1471-2288-12-132

Singh, G., Lingala, B. & Mithal, A. (2019) Gout and Hyperuricemia in the US: Prevalence and Trends. *Rheumatology*. 58(12), 2177–2180. doi:10.1093/rheumatology/kez196

Singh, J. A., Bharat, A., Khanna, D., Aquino-Beaton, C., Persselin, J. E., Duffy, E., Elashoff, D. & Khanna, P. P. (2016) Racial differences in health-related quality of life and functional ability in patients with gout. *Rheumatology*. 56(1), 103–112. doi:10.1093/rheumatology/kew356

Singh, J. A. & Cleveland, J. D. (2018a) Gout and the risk of incident atrial fibrillation in older adults: a study of US Medicare data. *Rheumatic & musculoskeletal diseases open*. 4(2), e000712. doi:10.1136/rmdopen-2018-000712

Singh, J. A. & Cleveland, J. D. (2018b) Gout and the risk of age-related macular degeneration in the elderly. *PLoS ONE*. 13(7), e0199562. doi:10.1371/journal.pone.0199562

Singh J.A. & Strand, V. (2008) Gout is associated with more comorbidities, poorer health-related quality of life and higher healthcare utilisation in US veterans. *Annals of the Rheumatic Diseases*. 67(9), 1310–1316. doi:10.1136/ard.2007.081604

Singh, J. A., Taylor, W. J., Simon, L. S., Khanna, P. P., Stamp, L. A., McQueen, F. M., Neogi, T., Gaffo, A. L. & Schumacher, H. R. (2011a) Patient-reported Outcome in Chronic Gout: A Report from OMERACT 10. *The Journal of Rheumatology*. 38(7), 1452–1457. doi:10.3899/jrheum.110271

Singh, J. A., Yang, S., Strand, V., Simon, L., Forsythe, A., Hamburger, S. & Chen, L. (2011b) Validation of pain and patient global scales in chronic gout: data from two randomised controlled trials. *Annals of the Rheumatic Diseases*. *70*(7), 1277–81. doi:10.1136/ard.2010.144022

Sivertsen, H., Bjørkløf, G. H., Engedal, K., Selbæk, G. & Helvik, A. S. (2015) Depression and quality of life in older persons: A review. *Dementia and Geriatric Cognitive Disorders*. 40(5–6), 311–339. doi:0.1159/000437299

Skevington, S, M., Lotfy, M. & O'connell, K, A. (2004) The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial A Report from the WHOQOL Group. *Quality of Life Research*. 13 (2), 299–310. doi:10.1023/B:QURE.0000018486.91360.00

Spaetgens, B., Van der linden, S. & Boonen, A. (2014) The gout assessment questionnaire 2.0: Cross-cultural translation into dutch, aspects of validity and linking to the international classification of functioning, disability and health. *Rheumatology*. 53(4), 678–685. doi:10.1093/rheumatology/ket423

Spaetgens, B., Wijnands, J. M. A., van Durme, C. & Boonen, A. (2015) Content and construct validity of the Rheumatic Diseases Comorbidity Index in patients with gout. *Rheumatology*. 54(9), 1659–1663. doi:10.1093/rheumatology/kev030

Spencer, E. A., Appleby, P. N., Davey, G. K. & Key, T. J. (2002) Validity of self-reported height and weight in 4808 EPIC-Oxford participants. *Public Health Nutrition*. *5*(4), 561–5. doi:10.1079/PHN2001322

Spencer, K., Carr, A. & Doherty, M. (2012) Patient and provider barriers to effective management of gout in general practice: A qualitative study. *Annals of the Rheumatic Diseases*. 71(9), 1490–1495. doi:10.1136/annrheumdis-2011-200801

Spitzer, R., Kroenke, K., Williams, J. & Lowe, B. (2006) A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*. 166(10), 1092–1097. doi:10.1001/archinte.166.10.1092

Spriensma, A. S., Eekhout, I., De Boer, M. R., Luime, J. J., De Jong, P. H., Bahçecitapar, M. K., Heymans, M.W. & Twisk, J. W. R. (2018). Analysing outcome variables with floor effects due to censoring: A simulation study with longitudinal trial data. *Epidemiology Biostatistics and Public Health*. 15(2), e12850. doi:10.2427/12850

Stamp, L. K., Taylor, W. J., Jones, P. B., Dockerty, J. L., Drake, J., Frampton, C. & Dalbeth, N. (2012) Starting dose is a risk factor for allopurinol hypersensitivity syndrome: A proposed safe

starting dose of allopurinol. *Arthritis and Rheumatism*. 64(8), 2529–2536. doi:10.1002/art.34488

StataCorp. (2013) *Stata multilevel mixed-effects reference manual: release 13*. College Station Texas, Stata Press.

Steiger, J.H. & Lind, J.C. (1980) Statistically-based tests for the number of common factors. Paper presented at the Annual Spring Meeting of the Psychometric Society in Iowa City. May 30, 1980

Stewart, S., Dalbeth, N., Vandal, A. C. & Rome, K. (2016) The first metatarsophalangeal joint in gout: a systematic review and meta-analysis. *BMC Musculoskeletal Disorders*. 17, 69. doi:10.1186/s12891-016-0919-9

Stewart, S., Rome, K.A.E., House, M. Horne, A., Doyle, A. J., Knight, J., Taylor, W. J. & Dalbeth, N. (2018) Predictors of activity limitation in people with gout: a prospective study. *Clinical Rheumatology*. 37 (8), 2213–2219. doi:10.1007/s10067-018-4110-6

Steyerberg, E.W., Moons, K.G.M, van der Windt, D.A. Hayden, J.A., Perel, P., Schroter, S., Riley, R.D., Hemingway, H. & Altman, G.D. (2013) Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. *PLoS Medicine*. 10(2), e1001381. doi:10.1371/journal.pmed.1001381

Stoel, R.D., van den Wittenboer, G. & Hox, J. (2004) Methodological Issues in the Application of the Latent Growth Model. In: Van Monfort, K., Oud, J. & Satorra, A. (eds.) *Recent developments in structural equation models*. Dordrecht, Kluwer Academic Publishers. 241-261.

Strand, V., Khanna, D., Singh, J.A., Forsythe, A. & Edwards, N.L. (2012) Improved health-related quality of life and physical function in patients with refractory chronic gout following treatment with pegloticase: Evidence from phase III randomized controlled trials. *Journal of Rheumatology*. 39 (7), 1450–1457. doi:10.3899/jrheum.111375

Strauss, V. Y., Jones, P. W., Kadam, U. T. & Jordan, K. P. (2014) Distinct trajectories of multimorbidity in primary care were identified using latent class growth analysis. *Journal of Clinical Epidemiology*. 67(10), 1163–1171. doi:10.1016/j.jclinepi.2014.06.003

Streeter, A. J., Lin, N. X., Crathorne, L., Haasova, M., Hyde, C., Melzer, D. & Henley, W. E. (2017) Adjusting for unmeasured confounding in nonrandomized longitudinal studies: a methodological review. *Journal of Clinical Epidemiology*. 87, 23–34. doi:10.1016/j.jclinepi.2017.04.022

Sultan, A., Mallen, C., Hayward, R., Muller, S., Whittle, R., Hotston, M. & Roddy, E. (2017) Gout and subsequent erectile dysfunction: A population-based cohort study from England. *Arthritis Research & Therapy*. 19, 123. doi:10.1186/s13075-017-1322-0

Szklo, M. & Nieto, F. J. (2014) *Epidemiology beyond the basics*. 3rd edition. Burlington Massachusetts, Jones & Bartlett Learning.

Taillefer, M., Dupuis, G., Roberge, M. & Le May, S. (2003) Health-Related Quality of Life Models: Systematic Review of the Literature. *Social Indicators Research*. 64(2), 293–323. doi:10.1023/A:1024740307643

Taylor, W. J., Colvine, K., Gregory, K., Collis, J., McQueen, F. M. & Dalbeth, N. (2008) The Health Assessment Questionnaire Disability Index is a valid measure of physical function in gout. *Clinical and Experimental Rheumatology*. *26*(4), 620–6. https://www.clinexprheumatol.org/abstract.asp?a=3414

Taylor, W. J., Fransen, J., Jansen, T. L., Dalbeth, N., Schumacher, H. R., Brown, M., Louthrenoo, W., Vazquez-Mellado, J., Eliseev, M., McCarthy, G., Stamp, L. K., Perez-Ruiz, F., Sivera, F., Ea, H. K., Gerritsen, M., Scire, C., Cavagna, L., Lin, C., Chou, Y. Y., Tausche, A. K., Vargas-Santos, A. B., Janssen, M., Chen, J. H., Slot, O., Cimmino, M. A., Uhlig, T. & Neogi, T. (2015) Study for updated gout classification criteria: Identification of features to classify gout. *Arthritis Care & Research*. 67(9), 1304–1315. doi:10.1002/acr.22585

Taylor, W. J., Shewchuk, R., Saag, K. G., Schumacher, H. R., Singh, J. A., Grainger, R., Edwards, N. L., Bardin, T., Waltrip, R. W., Simon, L. S. & Burgos-Vargas, R. (2009) Toward a Valid

Definition of Gout Flare: Results of Consensus Exercises Using Delphi Methodology and Cognitive Mapping. *Arthritis Care & Research*. 61(4), 535–543. doi:10.1002/art.24166

ten Klooster, P. M., Oude Voshaar, M. A. H., Taal, E. & van de Laar, M. A. F. J. (2010) Comparison of measures of functional disability in patients with gout. *Rheumatology*. 50(4), 709–713. doi:10.1093/rheumatology/keq387

ten Klooster, P. M., Oude Voshaar, M. A. H., Taal, E. & Van De Laar, M. A. F. J. (2011) Comparison of measures of functional disability in patients with gout. *Rheumatology. 50*(4), 709–713. doi:10.1093/rheumatology/keq387

ten Klooster, P. M., Vonkeman, H. E., Oude Voshaar, M. A. H., Bode, C. & Van De Laar, M. A. F. J. (2014) Experiences of gout-related disability from the patients' perspective: A mixed methods study. *Clinical Rheumatology*. 33(8), 1145–1154. doi:10.1007/s10067-013-2400-6

Teng, G. G., Nair, R. & Saag, K. G. (2006) Pathophysiology, clinical presentation and treatment of gout. *Drugs*. 66(12), 1547–63. doi:10.2165/00003495-200666120-00002

Ter Maaten, J. C., Voorburg, A., Heine, R. J., Ter Wee, P. M., Donker, A. J. M. & Gans, R. O. B. (1997) Renal Handling of Urate and Sodium during Acute Physiological Hyperinsulinaemia in Healthy Subjects. *Clinical Science*. 92(1), 51–58. doi: 10.1042/cs0920051

Testa, M. A. & Simonson, D. C. (1996) Assessment of Quality-of-life outcomes. *New England Journal of Medicine*. 334(13), 835–840. doi:10.1056/NEJM199603283341306

Trifirò, G., Morabito, P., Cavagna, L., Ferrajolo, C., Pecchioli, S., Simonetti, M., Bianchini, E., Medea, G., Cricelli, C., Caputi, A. P. & Mazzaglia, G. (2013) Epidemiology of gout and hyperuricaemia in Italy during the years 2005-2009: A nationwide population-based study. *Annals of the Rheumatic Diseases*. 72(5), 694–700. doi:10.1136/annrheumdis-2011-201254

Tucker, L. R. & Lewis, C. (1973) A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. *38*(1), 1–10. doi:10.1007/BF02291170

Twisk, J. & de Vente, W. (2002) Attrition in longitudinal studies: how to deal with missing data. *Journal of Clinical Epidemiology*. 55(4), 329–337. doi:10.1016/S0895-4356(01)00476-0

Twisk, J. & Hoekstra, T. (2012) Classifying developmental trajectories over time should be done with great caution: A comparison between methods. *Journal of Clinical Epidemiology*. *65*, 1078–1087. doi:10.1016/j.jclinepi.2012.04.010

Twisk, J. W. R. (2003) *Applied Longitudinal Data Analysis for Epidemiology A Practical Guide*. 1st edition. Cambridge, Cambridge University Press.

Twisk, J. W. R (2006) *Applied Multilevel Analysis A Practical Guide for Medical Researchers.* Cambridge, Cambridge University Press.

Twisk, J., De Boer, M., De Vente, W. & Heymans, M. (2013) Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis. *Journal Clinical Epidemiology*. 66(9), 1022–1028. doi:10.1016/j.jclinepi.2013.03.017

Van Belle, G., Fisher, L. D., Heagerty, P. J. & Lumley, T. (2004). *Biostatistics: A Methodology for the Health Sciences*. 2nd edition. Chichester, John Wiley & Sons.

van de Schoot, R., Sijbrandij, M., Winter, S. D., Depaoli, S. & Vermunt, J. K. (2017) The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies. *Structural Equation Modeling*. *24*(3), 451–467. doi:10.1080/10705511.2016.1247646

Van Den Broeck, J., Cunningham, S. A., Eeckels, R. & Herbst, K. (2005) Data cleaning: Detecting, diagnosing, and editing data abnormalities. *PLoS Medicine*. *2*(10), e267. doi:10.1371/journal.pmed.0020267

Van Echteld, I., Wechalekar, M. D., Schlesinger, N., Buchbinder, R. & Aletaha, D. (2014) Colchicine for acute gout. *Cochrane Database of Systematic Reviews*. 8, CD006190. doi:10.1002/14651858.CD006190.pub2

Van Groen, M.M., Ten Klooster, P.M., Taal. E., Van De Laar, M.A.F.J. & Glas, C.A.W. (2010) Application of the health assessment questionnaire disability index to various rheumatic diseases. *Quality of Life Research*. 19(9), 1255–1263. doi:10.1007/s11136-010-9690-9

Van Leeuwen, K. D. B., Bohnen, A. M., Jacobs, M. L., Van Der Lei, J., Janssens, H. J. E. M., Koffeman, A. R., Bindels, P. J. E. & Bierma-Zeinstra, S. M. A. (2018) Patient-reported gout attack frequency and allopurinol use in general practice in the Netherlands: A prospective observational cohort study protocol. *BMJ Open*. 8, e024335. doi:10.1136/bmjopen-2018-024335

Vandenbroucke, J. P., Von Elm, E., Altman, D. G., Gøtzsche, P.C., Mulrow, C. D., Pocock, S. J., Poole, C., Schlesselman, J. J. & Egger, M. (2007) Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. *PLoS Medicine*. 4 (10), e297. doi:10.1371/journal.pmed.0040297

Verkleij, S. P. J., Hoekstra, T., Rozendaal, J. H., Waarsing, J. H., Koes, B. W., Luijsterburg, P. A. J. & Bierma-Zenstra, S. M. (2012) Defining discriminative pain trajectories in hip osteoarthritis over a 2-year time period. *Annals of the Rheumatic Diseases*. 71(9), 1517–1523. doi:10.1136/annrheumdis-2011-200687

Wall, G. C., Koenigsfeld, C. F., Hegge, K. A. & Bottenberg, M. M. (2010) Adherence to treatment guidelines in two primary care populations with gout. *Rheumatology International*. 30, 749–753. doi:10.1007/s00296-009-1056-7

Wallace, B., Khanna, D., Aquino-Beaton, C., Singh, J. A., Duffy, E., Elashoff, D. & Khanna, P. P. (2016) Performance of Gout Impact Scale in a longitudinal observational study of patients with gout. *Rheumatology*. 55(6), 982–990. doi:10.1093/rheumatology/kew007

Wang, J. & Schaalje, G. B. (2009) Model selection for linear mixed models using predictive criteria. *Communications in Statistics: Simulation and Computation*. 38 (4), 788–801. doi:10.1080/03610910802645362

Wang, M. & Bodner, T. E. (2007) Growth Mixture Modeling: Identifying and predicting unobserved subpopulations with longitudinal data. *Organisational Research Methods*. 10(4), 635–657. doi:10.1177/1094428106289397

Wang, Y., Lopez, J. M. S., Bolge, S. C., Zhu, V. J. & Stang, P. E. (2016) Depression among people with type 2 diabetes mellitus, US National Health and Nutrition Examination Survey (NHANES), 2005-2012. *BMC Psychiatry*. 16, 88. doi:10.1186/s12888-016-0800-2

Ware, J. E. & Sherbourne, C. D. (1992) The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual framework. *Medical Care*. 30(6), 473–483.

Ware, J. E., Kosinski, M., Bayliss, M.S., McHorney, C.A., Rogers, W. H. & Raczek, A. (1995) Comparison of methods for the scoring and statistical analysis of SF36 health profile and summary measures: Summary of results from the medical outcomes study. *Medical care*. 33, AS264–AS279.

Wechalekar, M., Vinik, O., Schlesinger, N. & Buchbinder, R. (2013) Intra-articular glucocorticoids for acute gout. *Cochrane Database of Systematic Reviews*. 4, CD009920. doi: 10.1002/14651858.CD009920.pub2

West, T. B., Welch, K. B. & Galecki, A. T. (2015) *Linear Mixed Model A Practical Guide Using Statistical Software*. 2nd edition. Boca Raton Florida, CRC Press Taylor & Francis Group.

Wickrama, K., Lee, T. K., O'Neal, C. W. & Lorenz, F. (2016) *Higher-Order Growth Curves and Mixture Modeling with Mplus A Practical Guide*. New York, Routledge.

Wilson, I. B. & Cleary, P. D. (1995) Linking Clinical Variables with Health-Related Quality of Life. *Journal of American Medical Association*. 273(1), 59–65. doi:10.1001/jama.1995.03520250075037

Winnard, D., Wright, C., Taylor, W. J., Jackson, G., Te karu, L., Gow, P. J., Arroll, B., Thornley, S., Gribben, B. & Dalbeth, N. (2012) National prevalence of gout derived from administrative health data in Aotearoa New Zealand. *Rheumatology*. 51(5), 901–909. doi:10.1093/rheumatology/ker361

Wolf, E. J., Harrington, K. M., Clark, S. L. & Miller, M. W. (2013) Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *Educational and Psychological Measurement*. *73*(6), 913–934. doi:10.1177/0013164413495237

Wolff Smith, L. & Beretvas, S. N. (2009) Estimation of the Standardized Mean Difference for Repeated Measures Designs. *Journal of Modern Applied Statistical Methods*. 8 (2), 600–609. doi:10.22237/jmasm/1257035160

Wood, R., Fermer, S., Ramachandran, S., Baumgartner, S. & Morlock, R. (2016) Patients with gout treated with conventional urate-lowering therapy: Association with disease control, health-related quality of life, and work productivity. *Journal of Rheumatology*. 43 (10), 1897–1903. doi:10.3899/jrheum.151199

World Health Organisation (WHO). (2000) *Obesity: Preventing and Managing the Global Epidemic*. Geneva, World Health Organisation. Available from: https://www.who.int/nutrition/publications/obesity/WHO_TRS_894/en/ Accessed: March 2017

Wu, E. Q., Patel, P. A., Mody, R. R., Yu, A. P., Cahill, K. E., Tang, J. & Krishnan, E. (2009) Frequency, risk, and cost of gout-related episodes among the elderly: Does serum uric acid level matter? *Journal of Rheumatology*. 36(5), 1032–1040. doi:10.3899/jrheum.080487

Wu, W., West, S.G. & Taylor, A.B. (2009) Evaluating Model Fit for Growth Curve Models: Integration of Fit Indices From SEM and MLM Frameworks. *Psychological Methods*. 14 (3), 183–201. doi:10.1037/a0015858

Yan, T. & Curtin, R. (2010) The Relation Between Unit Nonresponse and Item Nonresponse: A Response Continuum Perspective. *International Journal of Public Opinion Research*. 22 (4), 535–551. doi:10.1093/ijpor/edq037

Ye, P., Yang, S., Zhang, W., Lv, Q., Cheng, Q., Mei, M., Luo, T., Liu, L., Chen, S. & Li, Q. (2013) Efficacy and Tolerability of Febuxostat in Hyperuricemic Patients With or Without Gout: A

Systematic Review and Meta-Analysis. *Clinical Therapeutics*. 35(2), 180–189. doi:10.1016/j.clinthera.2012.12.011

Yoon, J. Y., Brown, R. L., Bowers, B. J., Sharkey, S. S. & Horn, S. D. (2015) Longitudinal Psychological Outcomes of the Small-scale Nursing Home Model: a Latent Growth Curve Zero-inflated Poisson Model. *International psychogeriatrics*. *27*(6), 1009–1016. doi:10.1017/S1041610214002865

Yu, C. Y. (2002) Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. PhD Thesis. Available from: https://www.statmodel.com/download/Yudissertation.pdf Accessed: March 2018

Zhang, X. H., Li, S. C., Xie, F., Lo, N. N., Yang, K. Y., Yeo, S. J., Fong, K.Y. & Thumboo, J. (2012c) An exploratory study of response shift in health-related quality of life and utility assessment among patients with osteoarthritis undergoing total knee replacement surgery in a tertiary hospital in Singapore. *Value in Health*. 15(s1), 72–78. doi:10.1016/j.jval.2011.11.011

Zhang, Y., Chen, C., Choi, H., Chaisson, C., Hunter, D., Niu, J. & Neogi, T. (2012a) Purine-rich foods intake and recurrent gout attacks. *Annals of the Rheumatic Diseases*. 71(9), 1448–1453. doi:10.1136/annrheumdis-2011-201215

Zhang, Y., Hedo, R., Rivera, A., Rull, R., Richardson, S. & Tu, X. M. (2019) Post hoc power analysis: is it an informative and meaningful analysis? *General Psychiatry*. 32(4), e100069. doi:10.1136/gpsych-2019-100069

Zhang, Y., Neogi, T., Chen, C., Chaisson, C., Hunter, D. J. & Choi, H. (2014) Low-dose aspirin use and recurrent gout attacks. *Annals of the Rheumatic Diseases*. 73(2), 385–390. doi:10.1136/annrheumdis-2012-202589

Zhang, Y., Neogi, T., Chen, C., Chaisson, C., Hunter, D. J. & Choi, H. K. (2012b) Cherry consumption and decreased risk of recurrent gout attacks. *Arthritis and Rheumatism*. 64(12), 4004–4011. doi:10.1002/art.34677

Zhang, Y., Peloquin, C. E., Dubreuil, M., Roddy, E., Lu, N., Neogi, T. & Choi, H. K. (2015) Sleep apnea and the risk of incident gout: A population-based, body mass index-matched cohort study. *Arthritis and Rheumatology*. 67(12), 3298–3302. doi:10.1002/art.39330

Zhang, Y., Woods, R., Chaisson, C. E., Neogi, T., Niu, J., McAlindon, T. E. & Hunter, D. (2006) Alcohol Consumption as a Trigger of Recurrent Gout Attacks. *American Journal of Medicine*. *119*(9), 800.e13-800.e18. doi:10.1016/j.amjmed.2006.01.020

Zhou, M. & Carin, L. (2015) Negative binomial process count and mixture modeling. *Institute of Electrical and Electronics Engineers transactions on pattern analysis and machine intelligence*. *37*(2), 307–320. doi:10.1109/TPAMI.2013.211

Zhu, L. & Gonzalez, J. (2017) Modeling floor effects in standardized vocabulary test scores in a sample of low SES Hispanic preschool children under the multilevel structural equation modelling framework. *Frontiers in Psychology*. 8, 2146. doi:10.3389/fpsyg.2017.02146

Appendix 1: Studies which investigated HRQOL in people with gout reported in the systematic review by Chandratre et al (2013) and identified in this thesis in a subsequent review of the literature

Author	Publication year	Study period	Study design	Country	Setting / data source	Sample size	HRQOL outcome measure
Alvarez-Nemegyei et al	2005	1999	Cross-sectional	Mexico	Rheumatology Service and Primary care medical units	90	HAQ-DI
Colwell et al	2006	Not reported (NR)	Prospective	United States	Phase II program of Febuxostat	126	GAQ 1.0
Roddy, Zhang & Doherty	2007c	NR	Cross-sectional	United Kingdom	Two GP practices	13684	WHO-QOL
Alvarez-Hernandez et al	2008	NR	Prospective	Mexico	Eight rheumatology departments	206	HAQ-DI
Hirsch et al	2008	NR	Cross-sectional	United States	Family practice, internal medicine, and rheumatology clinics	308	GIS
Khanna et al	2008	NR	Cross-sectional	United States	General medicine and rheumatology clinics; Veterans Affairs Medical Centre, and a private rheumatology practice	80	SF-36 HAQ-DI EQ-5D
Singh & Strand	2008	1996-1998	Cross-sectional	United States	Veterans seen in Veterans Integrated Service Network	70334	SF-36
Taylor et al	2008	NR	Cross-sectional	New Zealand	Rheumatology outpatients	73	HAQ-DI

Alvarez-Hernandez et al	2009	NR	Prospective	Mexico	NR	49	HAQ-DI AIMS MOS-20
Becker et al	2009	NR	Prospective	United States	Academic and private rheumatology practices	110	SF-36 HAQ-DI
Lee et al	2009	NR	Cross-sectional	United States	Arthritis and primary care clinics Veteran Affairs health care systems and surrounding communities	371	SF-36
Hirsch et al	2010	NR	Cross-sectional	United States	Family practice, internal medicine, and rheumatology	308	GIS
Sarkin et al	2010	NR	Cross-sectional	United states	Family practice, internal medicine, and rheumatology	260	GIS
Van Groen et al	2010	2005-2008	Cross-sectional	Netherlands	Hospital rheumatology outpatient clinic	102	SF-36 HAQ-DI
Dalbeth et al	2011	NR	Prospective	New Zealand	Primary and secondary care settings	142	HAQ-II
Khanna et al	2011a	NR	Prospective	Spain	Academic centre-based gout clinic	99	SF-36
Khanna et al	2011b	NR	Prospective	United States	RCT assessing riloacept vs placebo	73	GIS
ten klooster et al	2011	2005-2008	Cross-sectional	Netherlands	Outpatient rheumatology clinic	102	HAQ-DI HAQ-II SF-36 PF10

Khanna et al	2012c	2010	Cross-sectional	France, Germany, United Kingdom	US and EU National Health and Wellness Surveys	1936	SF-12v2
Strand et al	2012	NR	Prospective	United States	RCT of pegloticase in refractory gout	212	SF-36 HAQ-DI
Dalbeth et al	2013	NR	Prospective	New Zealand	Primary and secondary care clinics	291	SF-36 HAQ-II
Scire et al	2013	June 2011-Jan 2012	Cross-sectional	Italy	Rheumatology clinics	446	SF-36 HAQ-DI
ten Klooster et al	2014	July - Dec 2012	Cross-sectional	Netherlands	Outpatient rheumatology clinics	34	HAQ-DI
Khanna et al	2015	April 2012-July 2012	Cross-sectional	United States	Health plan in the US Patients on healthcare plan were mailed survey	257	GIS SF-12
Spaetgens et al	2015	NR	Cross-sectional	Netherlands	University Hospital outpatient rheumatology clinic	122	GIS SF-36 HAQ-DI
Singh et al	2016	NR	Prospective cohort	United States	Primary care, rheumatology and other specialist clinics Veterans Affairs institutions	186	GIS SF-36 HAQ-DI
Wallace et al	2016	NR	Prospective	United States	Rheumatology or primary care	147	GIS SF-36 HAQ-DI
Wood et al	2016	NR	Cross-sectional	France, Germany, UK, United States	Rheumatology and primary care	1204	HAQ EQ-5D
Fu et al	2017	Nov 2015-Jan 2017	Cross-sectional	China	University hospital outpatients and inpatients	226	SF-36 HAQ-DI

Lopez Lopez et al	2017	2010-2014	Cross-sectional	Mexico	Secondary care	248	HAQ-DI
Chandratre et al	2018	2012	Cross-sectional	United Kingdom	Primary care	1184	GIS SF-36 PF10 HAQ-DI
La-Crette et al	2018	2015	Cross-sectional	United Kingdom	Academic rheumatology clinic	102	GIS
Stewart et al	2018	NR	Prospective	New Zealand	Primary and secondary care clinics	295	HAQ-II
Chincilla, Doherty & Aberhishek	2019	NR	Prospective	United Kingdom	RCT recruited from primary care	517	GIS
Edwards et al	2019	June 2006-Oct 2007	Cross-sectional	United States, Canada, Mexico	RCT of Pegloticase	212	SF-36 HAQ-DI
Lee et al	2019	NR	Cross-sectional	Singapore	Tertiary academic medical centre	267	GIS
Proudman et al	2019	2017	Cross-sectional	Australia	Population survey	2778	SF-12 v1

Appendix 2: Factors associated with disease-specific and generic HRQOL in people living with gout in published studies

Authors	Year	HRQOL outcome measure	Flare frequency	Gout severity◇	Current flare	Recent flare*	Number joints†	Disease duration	Tophi	Serum urate	Allopurinol	Colchicine	NSAIDS	Comorbidities	Pain	Depression	Anxiety	Age	Sex	Ethnicity	Obesity/BMI	Deprivation	Alcohol	Education	Marital status
Alvarez-Nemegyei et al	2005	HAQ-DI							✓				×	✓				×						×	
Roddy, Zhang & Doherty	2007	WHO-QOL								×	×			✓				✓	✓		✓				
Alvarez-Hernandez et al	2008	HAQ-DI					✓		✓						✓										
Hirsch et al	2008	GIS	✓	✓		✓				✓															
Alvarez-Hernandez et al	2009	HAQ-DI													✓	×	×								
Becker et al	2009	SF-36	✓				✓	×	✓	×				✓	✓										
		HAQ-DI	×				✓	×	×	×				✓	✓										
Lee et al	2009	SF-36	✓				✓							✓	✓			✓							
Hirsch et al	2010	GIS	✓	✓			✓		✓	✓					✓										
Sarkin et al	2010	GIS		✓																					
Dalbeth et al	2011	HAQ-II																✓							
Khanna et al	2011	SF-36	✓							✓								✓							
Khanna et al	2012	SF-12	✓						✓																
Dalbeth et al	2013	SF-36																		✓					
		HAQ-DI																		✓					
Scire et al	2013	SF-36	✓			✓	✓	✓	✓	×	×	✓	✓	✓							✓			✓	
		HAQ-DI	✓			✓	✓	✓	✓	×	×	✓	✓	✓							✓			✓	
Khanna et al	2015	GIS			✓																				
		SF-12			✓																				
Spaetgens et al	2015	GIS												×											
		SF-36												✓											
		HAQ-DI												✓											

Singh et al	2016	GIS																		✓					
		SF-36																		✓					
		HAQ-DI																		×					
Wallace et al	2016	GIS		✓		✓	✓	×	×	×				×				✓							
		SF-36					✓											✓							
		HAQ-DI		✓			✓																		
Fu et al	2017	SF-36 PCS	✓									✓		✓	✓	✓									
		SF-36 MCS													✓	✓	✓								
		HAQ-DI										✓			✓	✓									
Lopez Lopez et al	2017	HAQ-DI													✓										
Chandratre et al	2018	GIS	✓		✓		✓	✓			✓			✓	✓	✓	✓	✓	×	✓	×	✓	✓	×	×
		SF-36 PF10	✓		×		×	×			×			✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓
		HAQ-DI	✓		✓		✓	×			×			✓	✓	✓	✓	✓	✓	×	✓	✓	✓	×	×
La-Crette et al	2018	GIS		✓																					
Stewart et al	2018	HAQ-II	✓				✓			✓				✓	✓			✓			✓				
Chincilla,Doherty & Abhishek,	2019	GIS	✓																						
Edwards et al	2019	SF-36							✓																
		HAQ-DI							✓																
Lee et al	2019	GIS	✓		✓				×	✓				✓		✓	✓								
Proudman et al	2019	SF-12	✓								×														

✓ Association reported in the literature (however association may not be with all subscales of this HRQOL measure)
× A lack of association explicitly reported in the study
◊ Association with 'gout severity' or 'gout activity'
* Association with 'recent flares' or flare(s) in the previous three months
†Association with oligo/polyarticular gout, polyarticular gout, or the number of tender/swollen/painful joints

Appendix 3: Cover letter sent with baseline questionnaire

GP Practice Name
GP Practice Address
GP Practice Telephone Numbers

Patient's Name
Patient's Address
Patient's Postcode

Date *Study ID*

Dear (insert name),

The doctors in this practice are working with researchers in the Arthritis Research UK Primary Care Centre, at Keele University. We are writing to you to see if you would be willing to help us with a research project.

Researchers at Keele University are trying to find out about **gout**, to get a better understanding of this condition. Further details of the project are on the accompanying Participant Information Sheet.

You have been sent this letter because you have been to see your GP with gout or have taken medications for gout during the last two years. We hope that you will be able to spare a short amount of time to complete the enclosed questionnaire. It should take you no more than 30 minutes to fill in.

All of your answers will be dealt with in strict confidence. We can also assure you that whether or not you answer the questionnaire will not in any way affect the care you receive from this practice or elsewhere.

We would be very grateful if you would return the questionnaire in the envelope provided in the next two weeks. **You do not need a stamp.** A short while after this date, we will send a reminder to people whose questionnaire we have not received. If you would like to know more about this study, please contact **Priyanka Chandratre**, at Keele University on **01782 734721**.

We will be asking you if you would be willing to help with this research study in the future, so we will also ask your permission to contact you again at the end of the questionnaire. In addition, we will ask your permission for review of your medical records. Full details of this research study are provided in the enclosed Participant Information Sheet.

Thank you very much for your help with this research project.

Yours sincerely,

Name of GP(s)

Enc: Participant information sheet, gout questionnaire, pre-paid envelope

Appendix 4: Participant information sheet (PIS) for prospective cohort study

Arthritis Research UK
Providing answers today and tomorrow

Keele University

PRIMARY CARE SCIENCES
ARTHRITIS RESEARCH UK PRIMARY CARE CENTRE

PATIENT INFORMATION SHEET
REC Reference Number 12/NW/0297
Version 2.0, dated 08/05/12

The Gout Study

You are being invited to take part in a research study. Before you decide whether to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

What is the purpose of the study?

Your GP practice, together with Keele University, is carrying out a research study on gout. Gout is the most common cause of inflamed joints in adults and can recur from time to time. Despite this, little is known about the way gout can affect peoples' quality of life and how to identify and treat those who may be at risk of having a worse outcome than others. We are trying to find out more about how gout affects people in the community and how it changes over the time.

Why have I been invited?

You were selected because you have been to see your GP with gout or you have taken medication for gout during the last two years.

Do I have to take part?

Whether or not you take part in this research is up to you. If you do decide to take part, you are free to withdraw at any time without giving a reason. A decision to withdraw, or a decision not to take part, will not affect your right to access health services at your practice or elsewhere.

How long will it take?

Taking part in this study means that you are asked to complete the enclosed questionnaire. We will also send you another similar

questionnaire in 6, 12, 24 and 36 months' time. We think it should take up to 30 minutes to complete each questionnaire.

Future contact

In the future, we may contact you again to ask you further questions about gout. We ask for your permission to contact you again on the last page of the questionnaire. If you agree to be contacted again, this does not mean that you must take part in future; you are only agreeing to be contacted again.

What are the possible benefits of taking part?

Although there are no immediate benefits to you as a patient, we hope that the insight we gain from this research will help patients in the future. Your participation will help us to learn more about gout and how to best treat it in general practice.

What are the possible risks of taking part?

The nature of the questionnaire is not meant to be distressing in any way. However if the questionnaires lead to distress, unpleasant memories or thoughts, we would encourage you to contact your General Practitioner. You may also wish to contact an independent mental health support group, which does not require referral from a doctor or a nurse. All calls are free (call back also available), confidential and support is provided by trained staff. The phone numbers of these support groups will also be provided at the end of each questionnaire.

Mental Health Helpline Staffordshire (Brighter Futures) 0808 800 2234

Mental Health Helpline Shropshire 0800 195 1700

Mental Health Helpline Wolverhampton 0800 387034

Will my taking part in this study be kept confidential?

The answers you give in the questionnaire will be dealt with in **strictest confidence**. Each person who responds to the questionnaire will be given a code number, so the data from the study will not have any identifiable names and addresses, and cannot be traced back to you. On this basis, the data may be used in other research studies.

How long will the answers to the study questionnaires be stored for?

The questionnaires will be stored without identifiable names and addresses for twenty years in accordance with the Medical Research

Council guidelines. Beyond this date records will be maintained if the study is still on-going. The questionnaires will be stored in a secure place. Any identifiable personal information such as your name and address will however be destroyed at the end of the study period. This will ensure that personal data will not be stored for longer than is necessary (Data Protection Act 1998).

Medical record review
Another part of this study is to find out what other factors related to gout, such as your medications and other health problems may influence your quality of life. We can do this by reviewing your medical records, and we ask your permission to do this on the last page of the questionnaire. When reviewing medical records, your name will not be used so that you cannot be identified personally. All information will be held in strictest confidence.

What will happen if I don't want to carry on with this study?
You can withdraw from this study at any stage by contacting **Priyanka Chandratre**, the Gout Study Co-ordinator on **01782 734721**. Withdrawing means that we would no longer contact you directly, but we would still keep and use the information you have provided up to the point of your withdrawal. If you contact us to withdraw from the study, and you have consented to medical record review, we will check whether you also want us to stop reviewing your medical records.

What will happen to the results of the research study?
Because this is a large study, the results will not be available for about three years, and will then be published in medical journals and reports. The main findings from the study will be displayed on a poster in your practice. If you would like any other information after seeing this poster we will be happy to help.

Who is funding and organising the research?
The research is funded and organised by the Arthritis Research UK Primary Care Centre at Keele University.

Who has reviewed the study?
The Liverpool East Research Ethics Committee has reviewed this study (Research Ethics Committee Reference Number: 12/NW/0297).

Contact for further information
If you have any questions, or would like further information, about this study please contact **Priyanka Chandratre**, the Gout Study Co-

ordinator on **01782 734721**. If you have any questions or concerns about taking part in this research you can also contact the Patient Advice and Liaison Service (PALS). Your local PALS office phone number for NHS Stoke-on-Trent is 0800 783 2865, Wolverhampton PCT is 01902 445378, NHS Telford and Wrekin is 01952 580478, NHS Shropshire county is 01952 580474, NHS South Staffordshire is 01543 465106 and for NHS North Staffordshire is 0800 030 4563.

Thank you for taking time to read this information leaflet.

Appendix 5: Baseline questionnaire and consent forms

Arthritis Research UK
Providing answers today and tomorrow

Keele University

PRIMARY CARE SCIENCES
ARTHRITIS RESEARCH UK PRIMARY CARE CENTRE

Gout Study Questionnaire

Baseline Questionnaire Booklet

REC Reference Number: 12/NW/0297

INSTRUCTIONS FOR THIS QUESTIONNAIRE

Please answer **all** the questions.

The questions can be answered by putting a cross in a box like this: ☒

or circling a number like this: 3 (4) 5 6

Please write in BLOCK CAPITALS where appropriate. **Please complete the consent form on page 22 if you agree to take part in this study, then complete the questionnaire.**

When you have finished please check that you have answered all of the questions and then return the questionnaire in the envelope enclosed. **You do not need a stamp**. Please return the questionnaire in the next two weeks.

The answers you give in the questionnaire will be treated in the strictest confidence.

Whether you take part in this research or not, your right to use health services at your practice or elsewhere will not be affected.

Details about this project are available in the information sheet enclosed. If you would like further information please contact Priyanka Chandratre on **01782 734721**

THANK YOU

SECTION A: ABOUT GOUT

1. How many attacks of gout have you had in the last 12 months?
   (Please put a cross in one box only)

   0..... ☐ 2..... ☐ 4..... ☐

   1..... ☐ 3..... ☐ 5 or more.. ☐

2. How old were you when you were first diagnosed with gout?

   Age ☐☐ Years

3. Are you having an attack of gout at the present?

   Yes ☐ No ☐

4. Have you ever had gout in more than one joint at the same time?

   Yes ☐ No ☐

5. Do you currently take a tablet called allopurinol for gout?

   Yes ☐ No ☐

   If yes, please indicate the dose below

   50 mg..... ☐
   100 mg..... ☐
   200 mg..... ☐
   300 mg..... ☐
   400 mg..... ☐
   500 mg..... ☐
   600 mg..... ☐
   700 mg..... ☐
   800 mg..... ☐
   900 mg..... ☐
   Don't know..... ☐
   Other (please specify)..... ☐

SECTION B: HOW GOUT AFFECTS YOUR LIFE

1. Please indicate how much you agree or disagree with the statements below **(Please put a cross in one box only for each statement).**

	Strongly disagree	Disagree	Uncertain	Agree	Strongly agree
a. I am worried that I will have a gout attack within the next year	☐	☐	☐	☐	☐
b. I am afraid that my gout will get worse over time	☐	☐	☐	☐	☐
c. I worry that I will not be able to continue to enjoy my leisure activities as a result of my gout	☐	☐	☐	☐	☐
d. I feel anxious that my gout will interfere with my future activities	☐	☐	☐	☐	☐
e. I am bothered by the side effects from my gout medications	☐	☐	☐	☐	☐
f. I am mad or angry when I experience a gout attack	☐	☐	☐	☐	☐
g. It is difficult to plan ahead for events or activities because I may have a gout attack	☐	☐	☐	☐	☐
h. I feel depressed when I get a gout attack	☐	☐	☐	☐	☐
i. My current medications are effective at treating a gout attack if I get one	☐	☐	☐	☐	☐

	Strongly disagree	Disagree	Uncertain	Agree	Strongly agree
j. I miss planned or important activities when I have a gout attack	☐	☐	☐	☐	☐
k. I worry about the long term effects of my gout medications	☐	☐	☐	☐	☐
l. My current medications do not work well to prevent gout attacks from happening	☐	☐	☐	☐	☐
m. I have control over my gout	☐	☐	☐	☐	☐

2. **During your last gout attack**, how much of the time did you experience the following?
(Please put a cross in one box only for each statement).

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
a. Miss work because of gout symptoms?	☐	☐	☐	☐	☐
b. Have difficulty working because of gout symptoms?	☐	☐	☐	☐	☐
c. Have difficulty with recreational or social activities because of your gout symptoms?	☐	☐	☐	☐	☐
d. Have difficulty with self care such as bathing, feeding , dressing yourself because of gout symptoms?	☐	☐	☐	☐	☐

3. **During your last gout attack**, how much did your symptoms interfere with the following things?
(Please put a cross in one box only for each statement).

	Not a bit	A little bit	Moderately	Quite a bit	Extremely
a. Your mood?	☐	☐	☐	☐	☐
b. Your ability to move about?	☐	☐	☐	☐	☐
c. Your sleep?	☐	☐	☐	☐	☐
d. Your normal work? (including both work outside the home and housework)	☐	☐	☐	☐	☐
e. Your recreational activities?	☐	☐	☐	☐	☐
f. Your enjoyment of life?	☐	☐	☐	☐	☐
g. Your ability to do what you want to do?	☐	☐	☐	☐	☐

4. Please indicate how much you agree or disagree with the statements below
(Please put a cross in one box only for each statement).

	Strongly disagree	Disagree	Uncertain	Agree	Strongly agree
a. There is a lot I can do to control my gout	☐	☐	☐	☐	☐
b. What I will do will affect whether my gout gets better or worse	☐	☐	☐	☐	☐
c. Treatments are effective in controlling gout	☐	☐	☐	☐	☐
d. Gout is a serious condition	☐	☐	☐	☐	☐

SECTION C: ABOUT YOUR GENERAL HEALTH

We are interested in your general health. **Please answer every question.** Some questions may look similar to others but each one is different. Please take the time to read and answer each question carefully by placing a cross in the box of your choice.

1. The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?
   (Please put a cross in one box on each line)

	Yes, limited a lot	**Yes, limited a little**	**No, not limited at all**
a. **Vigorous activities,** such as running, lifting heavy objects, participating in strenuous sports	☐	☐	☐
b. **Moderate activities,** such as moving a table, pushing a vacuum cleaner, bowling or playing golf	☐	☐	☐
c. Lifting or carrying groceries	☐	☐	☐
d. Climbing **several** flights of stairs	☐	☐	☐
e. Climbing **one** flight of stairs	☐	☐	☐
f. Bending, kneeling or stooping	☐	☐	☐
g. Walking **more than a mile**	☐	☐	☐
h. Walking **half a mile**	☐	☐	☐
i. Walking **one hundred yards**	☐	☐	☐
j. Bathing and dressing yourself	☐	☐	☐

2. Please place a cross in the box which best describes your abilities over the **past one week**.

	Without any difficulty	**With some difficulty**	**With much difficulty**	**Unable to do**
a. Dress yourself including tying shoe-laces and doing buttons?	☐	☐	☐	☐
b. Shampoo your hair?	☐	☐	☐	☐
c. Stand up from a chair?	☐	☐	☐	☐
d. Get in and out of bed?	☐	☐	☐	☐
e. Cut your meat?	☐	☐	☐	☐
f. Open a milk carton?	☐	☐	☐	☐
g. Lift a full glass or cup to your mouth?	☐	☐	☐	☐
h. Walk outdoors on flat ground?	☐	☐	☐	☐
i. Climb up 5 steps?	☐	☐	☐	☐
j. Wash and dry your entire body?	☐	☐	☐	☐
k. Take a tub bath?	☐	☐	☐	☐
l. Get on and off the toilet?	☐	☐	☐	☐
m. Reach and get a 5 pound object such as a bag of sugar from above your head?	☐	☐	☐	☐

	Without any difficulty	With some difficulty	With much difficulty	Unable to do
n. Open car doors?	☐	☐	☐	☐
o. Open jars that have been previously opened?	☐	☐	☐	☐
p. Bend down and pick up clothing from the floor?	☐	☐	☐	☐
q. Turn taps on and off?	☐	☐	☐	☐
r. Run errands and shop?	☐	☐	☐	☐
s. Get in and out of a car?	☐	☐	☐	☐
t. Do chores such as vacuuming or yard work?	☐	☐	☐	☐

3. Do you use any aids or devices for any of the above activities?
(Please put a cross in as many boxes as apply)

a. Raised toilet seat..... ☐
b. Devices used for dressing (button, hook, zipper pull, shoe horn etc.)..... ☐
c. Bathtub bar..... ☐
d. Special or built-up chair.... ☐
e. Long-handled appliances for reach..... ☐
f. Built-up or special utensils..... ☐
g. Bathtub seat..... ☐
h. Cane..... ☐
i. Long-handled appliances in bathroom..... ☐
j. Walker..... ☐
k. Jar opener (for jars previously opened)..... ☐
l. Crutches..... ☐
m. Wheelchair..... ☐
n. Other (please specify)..... ☐

4. Do you receive any help from another person for;
(Please put a cross in as many boxes as apply)

a. Hygiene..... ☐
b. Dressing and Grooming..... ☐
c. Gripping and opening things..... ☐
d. Arising..... ☐
e. Reach..... ☐
f. Eating..... ☐
g. Errands and chores... ☐
h. Walking..... ☐

5. How much pain have you had in the **past one week**? On a scale of 0 to 10 (where 0 represents "no pain" and 10 represents "pain as bad as can be"), please circle the number below.

|_____|_____|_____|_____|_____|_____|_____|_____|_____|_____|_____|
0 1 2 3 4 5 6 7 8 9 10

No pain Pain as bad as can be

6. Please rate how well you are doing on a scale of 0 to 10 (0 represents "very well" and 10 represents "very poor" health).
Please circle the number below.

|_____|_____|_____|_____|_____|_____|_____|_____|_____|_____|_____|
0 1 2 3 4 5 6 7 8 9 10

Very well Very poor health

7. Have you ever been diagnosed as having or been treated for the following?
(Please put a cross in as many boxes as apply)

a. Diabetes..... ☐
b. Stroke..... ☐
c. High blood pressure..... ☐
d. TIA or mini stroke..... ☐
e. High levels of cholesterol, fats or lipids in your blood..... ☐
f. Kidney failure... ☐
g. Heart attack..... ☐
h. Kidney stones.. ☐
i. Angina..... ☐

SECTION D: ABOUT HOW YOU FEEL

1. The next set of questions are about how you have felt over the last 2 weeks. Please read each item and put a cross in the box that comes closest to how you have been feeling in the **past 2 weeks**.

	Not at all	Several days	More than half the days	Nearly every day
a. Little interest in doing things	☐	☐	☐	☐
b. Feeling down, depressed, or hopeless	☐	☐	☐	☐
c. Trouble falling/staying asleep, sleeping too much	☐	☐	☐	☐
d. Feeling tired or having little energy	☐	☐	☐	☐
e. Poor appetite or overeating	☐	☐	☐	☐
f. Feeling bad about yourself, or that you are a failure or have let yourself or your family down	☐	☐	☐	☐
g. Trouble concentrating on things, such as reading the newspaper or watching TV	☐	☐	☐	☐
h. Moving or speaking so slowly that other people could have noticed. Or the opposite - being so fidgety or restless that you have been moving around a lot more than usual.	☐	☐	☐	☐
i. Thoughts that you would be better off dead or of hurting yourself in some way	☐	☐	☐	☐

2. If you have been bothered by any of the nine problems above, please answer the following:

How difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?

(Please put a cross in one box only)

Not difficult at all	Somewhat difficult	Very difficult	Extremely difficult
☐	☐	☐	☐

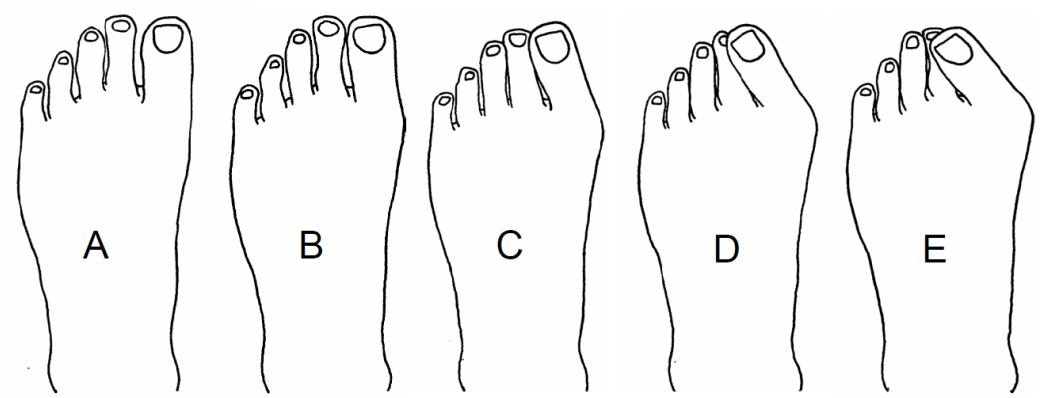
3. **Over the last 2 weeks**, how often have you been bothered by any of the following problems?
(Please put a cross in one box only for each statement)

	Not at all	**Several days**	**More than half the days**	**Nearly every day**
a. Feeling afraid that something awful may happen?	☐	☐	☐	☐
b. Worrying too much about different things?	☐	☐	☐	☐
c. Becoming easily annoyed or irritable?	☐	☐	☐	☐
d. Feeling nervous, anxious or on edge?	☐	☐	☐	☐
e. Not being able to stop or control worrying?	☐	☐	☐	☐
f. Trouble relaxing?	☐	☐	☐	☐
g. Being so restless that it is hard to sit still?	☐	☐	☐	☐

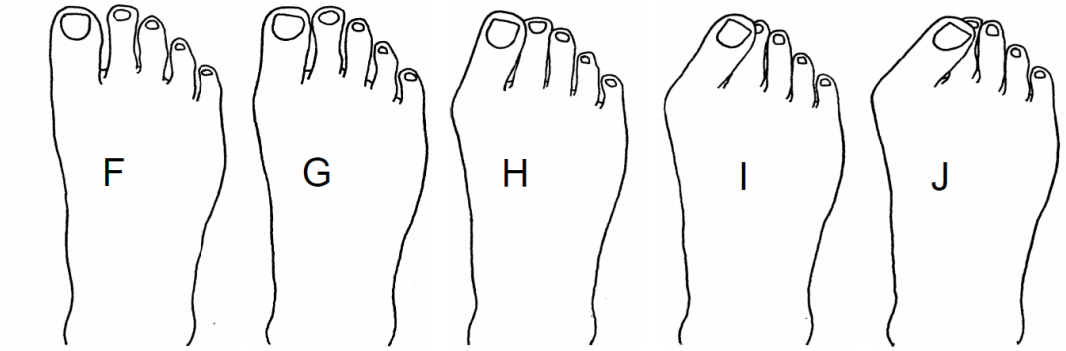
SECTION E: FOOT AND OTHER JOINT PROBLEMS

Part 1 - Feet

We are interested in whether your big toes are straight or angled sideways because this might be related to your ability to move around. First, please look at your left big toe whilst standing **without shoes and socks on**. Ignore the positioning and the gaps between your other toes and try to focus **only** on your big toe. Select from the first set of pictures below labelled from A to E which one best shows the angle of your left big toe. Please **circle** the letter of that picture.



Now do the same for your right big toe joint using the set of pictures below labelled from F to J. Again please **circle** the letter of the picture that best shows the angle of your right big toe.



Part 2: Pain and discomfort in the feet

1. In the **past month** have you had pain **or** aching **or** stiffness in your feet?

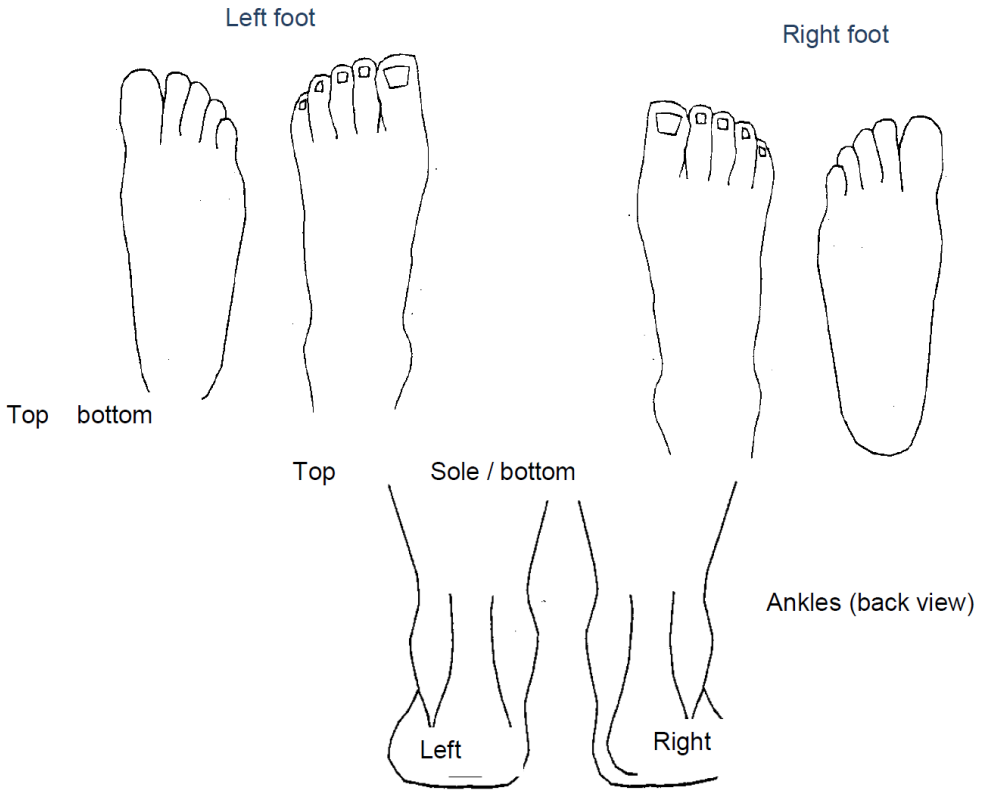
No days	Few days	Some days	Most days	All days
☐	☐	☐	☐	☐

If 'No days' please continue with **question 4 on page 16**

2. This question is about any **recent pain** you have had in your **feet**. In the **past month**, have you had <u>any</u> ache or pain that has lasted for **one day or longer** in your feet? Please **do not** include pain due to feverish illness such as flu.

Yes..... ☐ → ***Please shade*** *in the diagrams below any pain you have had in your feet in the last* ***month*** *that has lasted* ***one day or longer***

No..... ☐



3. Below are some statements about problems related to pain in the feet. For each statement indicate if this has applied to you **during the past month.**
(Please tick only one box for each statement).

	None of the time	On some days	On most / every day(s)
Because of pain in my feet:			
a. I avoid walking outside at all	☐	☐	☐
b. I avoid walking long distances	☐	☐	☐
c. I don't walk in a normal way	☐	☐	☐
d. I walk slowly	☐	☐	☐
e. I have to stop and rest my feet	☐	☐	☐
f. I avoid hard or rough surfaces when possible	☐	☐	☐
g. I avoid standing for a long time	☐	☐	☐
h. I catch the bus or use the car more often	☐	☐	☐
i. I need help with housework/ shopping	☐	☐	☐
j. I still do everything but with more pain or discomfort	☐	☐	☐
k. I get irritable when my feet hurt	☐	☐	☐
l. I feel self conscious about my feet	☐	☐	☐
m. I get self conscious about the shoes I have to wear	☐	☐	☐
n. I have constant pain in my feet	☐	☐	☐
o. My feet are worse in the morning	☐	☐	☐
p. My feet are more painful in the evening	☐	☐	☐
q. I get shooting pains in my feet	☐	☐	☐

4. Have you consulted your GP (family doctor) in the past 12 months because of problems with your foot or feet?

   (Please put a cross in one box only)

   Yes ☐ No ☐

5. Which of the following services have you used in the **past 12 months** because of problems with your foot or feet? For **each service** you have used please put a cross to show whether the NHS provided this, or if you had private treatment. If you have used both NHS **and** private services please cross both boxes. For any service you have not used please leave blank.

	NHS	**Private**
a. Physiotherapy	☐	☐
b. Podiatry	☐	☐
c. Chiropody	☐	☐

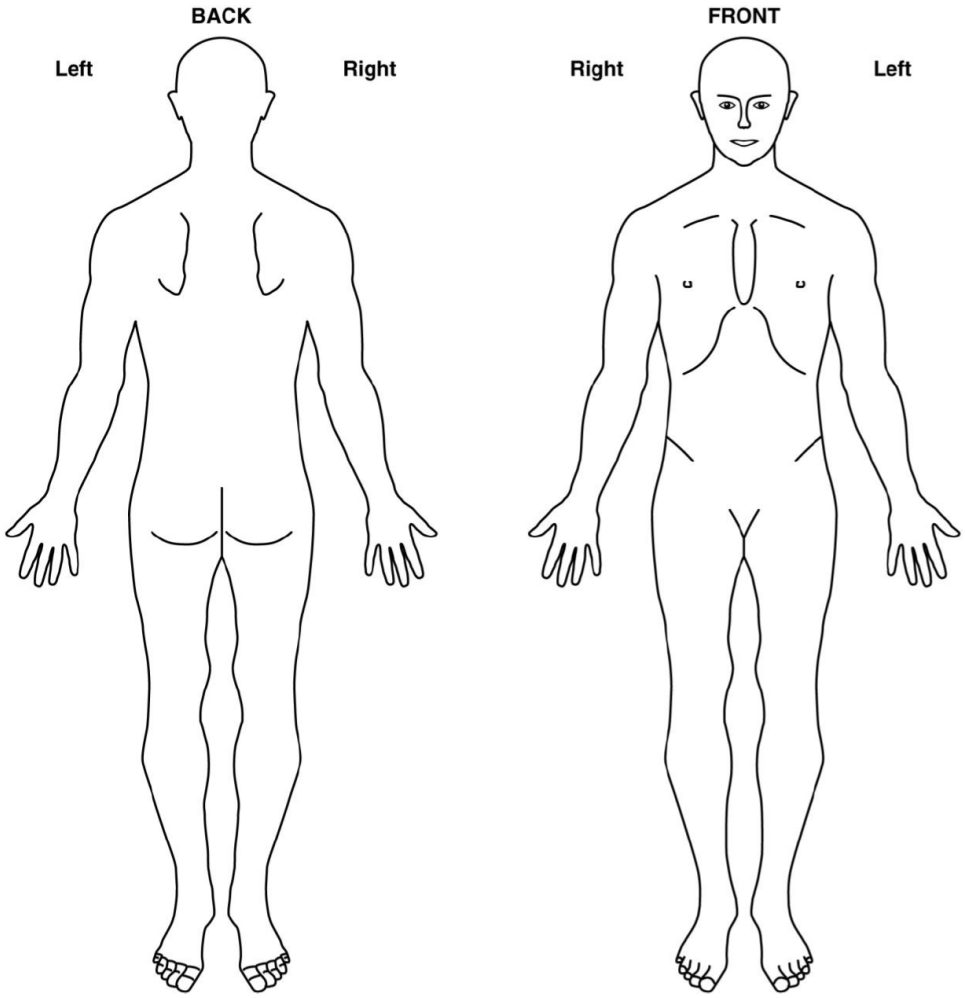
Part 3 – Body Chart

This question is about **recent pain** you may have had in **any part of your body**. By pain we also mean ache or discomfort or stiffness. Please **do not** include pain due to a feverish illness such as flu. If you are a woman please **do not** include pain related to your monthly period.

In the past 4 weeks**, have you had pain that has lasted for** one day or longer **in** any part of your body**?** ***(Please put a cross in one box only)***

Yes…... ☐ → **Please shade** in the diagram below **any pain** that has lasted for **one day or longer in the past 4 weeks**

No……. ☐ → Please continue with **section F on page 18**



SECTION F: HOW GOUT AFFECTS YOUR WORK

THIS SECTION ASKS SOME GENERAL QUESTIONS ABOUT YOUR WORK

1. Which of the following best describes your **current** situation?

(Please put a cross in one box only)

Working full-time in a paid job..... ☐

Working part-time in a paid job..... ☐

Employed, but currently off sick for 6 months or less..... ☐

Looking after the home / children..... ☐

Not working, for more than 6 months due to joint problems..... ☐

Not working, for more than 6 months for other reasons..... ☐

Fully retired..... ☐

Early retirement due to joint / back problems..... ☐

Early retirement for other reasons..... ☐

Student..... ☐

Other ..... ☐

If 'other', please write in your current situation below:-

.....

2. Have you **taken time off work** during the last 6 months because of gout?

(Please put a cross in one box only)

Yes ☐ No ☐

3. Are you currently

(Please put a cross in one box only)

a. Doing your usual job? ☐ *Please go to* ***Section G on page 20.***

b. Working fewer hours? ☐

c. Doing lighter duties? ☐

d. On paid sick leave? ☐

e. On unpaid leave? ☐

If you have answered b) to e), please answer question 4.

4. Yes ☐ No ☐ If you are not doing your usual job, is this because of joint problems?

SECTION G: ABOUT YOURSELF

THIS SECTION ASKS SOME GENERAL QUESTIONS ABOUT YOU

1. What is your date of birth?

☐☐/☐☐/☐☐

(E.g. – if you were born on the 5th of June 1936, this would be entered as 05/06/36)

2. Are you

Male ☐ Female ☐

3. What is your relationship status

(Please place a cross in one box only)

a. Married..... ☐ b. Widowed..... ☐

c. Co- habiting..... ☐ d. Divorced..... ☐

e. Separated..... ☐ f. Single..... ☐

4. Did you go on from school to full-time education or university?

Yes ☐ **If yes**, what age did you finish full-time education? ☐☐ years

No ☐

5. Is your ethnic origin? **(Please put a cross in one box only)**

a. White UK/European..... ☐ b. Asian..... ☐

c. Afro Caribbean..... ☐ d. African..... ☐

e. Chinese..... ☐ f. Other..... ☐

6. What is your height?

☐ Feet ☐☐ inches OR ☐☐☐ cms

7. What is your weight?

☐☐ Stones ☐☐ lbs OR ☐☐☐ kgs

8. About how often do you drink alcohol?

(Put a cross in one box only)

a. Daily or almost daily..... ☐
b. 3 to 4 times a week..... ☐
c. Once or twice a week..... ☐
d. 1 to 3 times a month..... ☐
e. Special occasions only..... ☐
f. Never..... ☐

9. In an **average week** how many

	Number
a. Small glasses (175 ml) of **wine** do you drink (there are roughly 6 glasses per bottle)?.....	☐☐
b. Pints of **beer** do you drink (includes bitter, lager, stout and ale)?..	☐☐
c. Measures of **spirits** do you drink (includes Whiskey)?.....	☐☐

Study ID

SECTION H: CONSENT FORM

Thank you for completing this questionnaire
Please ensure that you have read the enclosed information sheet that explains about the study and other similar questionnaires that will be sent to you in 6, 12, 24 and 36 months time. Please read and complete the following consent form, and then sign below.

Consent form
I confirm that I have read and understood the study information sheet and am willing to take part in the study. I understand that I can withdraw from the study at any time, and that this will not affect the care I receive in any way.

Please answer each statement by putting a cross in the box on each line

	Yes	No
I give my permission for my medical records to be reviewed……	☐	☐
I am happy to be contacted again (this does not mean that you must take part in future - you are just agreeing to be contacted again)……	☐	☐

Signed _____ Date _____

Please print your name and address _____

For office use only:

Even if you would prefer us not to review your medical records or contact you again about the study, the answers you have given in this questionnaire will still be very important to us. Please return your questionnaire in the **FREEPOST (no stamp needed)** envelope provided. Thank you for your help with this research project.

Appendix 6: Two week non-response reminder postcard sent after baseline questionnaire

GP Practice Name
GP Practice Address
GP Practice Telephone Numbers

Patient's Name
Patient's Address
Patient's Postcode

Date *Study ID*

The Gout Study

We are writing to remind you of a questionnaire that we recently sent to you. We are still interested in your response. We know that you may be busy, but it would be very helpful to us if you would take the time to complete the questionnaire and return it to us in the envelope that we previously provided. **You do not need a stamp**. We would be grateful if you could return the questionnaire in the next two weeks**. Your answers are strictly confidential**. If you have returned the questionnaire in the last few days, please ignore this postcard and we apologise for troubling you again. This study is being carried out by researchers from Keele University. If you have any questions about the questionnaire please feel free to contact **Priyanka Chandratre, the Gout Study Co-ordinator** on **01782 734721**.

Thank you very much for your help with this research. It is greatly appreciated.

Names of GP(s)

The Gout Study – Two week reminder postcard sent after baseline questionnaire, version 1.0; dated 15/02/12

Appendix 7: Four week non-response reminder cover letter sent with repeat questionnaire

GP Practice Name
GP Practice Address
GP Practice Telephone Numbers

Patient's Name
Patient's Address
Patient's Postcode

Date *Study ID*

Dear (insert name),

We are writing to remind you of the study that the doctors in this practice are carrying out with researchers from the Arthritis Research UK Primary Care Centre, at Keele University.

So far the researchers don't seem to have received a reply from you. We are therefore sending you a second questionnaire in case you mislaid the first one. We would be very grateful if you could spare a few minutes of your time to complete this questionnaire.

Your response is very important to us **– even if you are no longer having problems with your gout.** Your answers will be treated in the strictest confidence.

We would be grateful if you would fill in the questionnaire within the next two weeks and send it back in the envelope provided **(no stamp is needed)**. If you have returned the questionnaire within the last few days then please ignore this reminder, and we apologise for troubling you.

Your participation is voluntary and it is up to you whether you take part or not. Whether or not you return the questionnaire will not affect your health care in any way.

If you have any questions about the questionnaire please feel free to phone Priyanka Chandratre on **01782 734721.**

Thank you very much for your help with this research project.

Yours sincerely,

Names of GP(s)

Enc: Gout questionnaire, Patient Information Sheet, Pre-Paid Envelope

Appendix 8: Ethical approval letter for prospective cohort study

Revised Approved docs. Table 2012.06.27

NHS
Health Research Authority

NRES Committee North West - Liverpool East
HRA NRES Centre North West
Barlow House
3rd Floor
4 Minshull Street
Manchester
M1 3DZ

Telephone: 0161 625 7832
Facsimile: 0161 625 7299

30 April 2012

Dr Edward Roddy
Clinical Senior Lecturer in Rheumatology/ Honorary Consultant Rheumatologist
Keele University
Arthritis Research UK Primary Care Centre
Keele University
Staffordshire
ST5 5BG

Dear Dr Roddy

Study title:	**Prospective observational cohort study of health related quality of life and chronic foot problems and their determinants in gout.**
REC reference:	**12/NW/0297**
Protocol number:	**Protocol Number 1.0**

The Research Ethics Committee reviewed the above application at the meeting held on 19 April 2012. Thank you for attending with Dr Priyanka Chandratre to discuss the study.

Ethical opinion

The Committee welcomed you both to the meeting and thanked you for what they had found to be a very well-written study.

The Committee firstly asked why this study had not been submitted through the Proportionate Review service. Neither of you knew the reason.

The Committee had noted that in the Participant Information Sheet it was stated that questionnaires would be kept for 20 years and suggested that they be destroyed at the end of the study. You both agreed.

The Committee felt that there was a possibility that participants could be distressed and asked how this would be managed. Dr Chandratre explained that it is not intended that the participants would be distressed but if they were she would use her training to manage the situation. She would offer the participant the chance to take a break and step outside the room. She would ask if they still wanted their data to be included and would remove it if not. If the distress was longer term she would suggest they contact their GP. You emphasised that the questionnaire is not particularly likely to distress the participant.

The Committee made reference to question 1.i. in section D of the Gout Study Questionnaire which asks how often a person has had *'thoughts that you would be better off dead or of*

hurting yourself in some way.' The Committee felt that questions such as this would potentially distress participants. You explained that the questionnaire is validated and this is just a single question.

The Committee suggested that questions such as this may potentially trigger pre-existing distress. It was suggested that contact details for an alternative source of support such as a support group or charity were given. You both agreed.

The Committee asked about the GP practices to be involved in the study and how they would be notified as they had not seen a letter included with the application. You explained that the GP practices are already signed up to the research network.

The Committee asked if they would be including participants on the new gout drug Febuxostat. You explained that they would not. The prescription rate for this drug is very low.

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Ethical review of research sites

NHS Sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations

The Committee gave a favourable opinion of the application (with additional conditions as follows):

- Written clarification that questionnaires and tapes would be destroyed at the end of the study and not after 20 years should be provided. This should be amended on page 2 of the Patient Information Sheets (questionnaires and interviews).
- A section should be included in the Patient Information Sheet (questionnaires) about the risks of participating in the study. Information should be included in this section about the possibility of becoming distressed during the study and what should be done in this situation. Information should be included about support groups or charities that could be contacted if available.
- The location of the interview should be given in the Patient Information Sheet (interviews).
- It should be made clear in the Patient Information Sheet (interviews) that if distressed the participant could leave the interview.
- The contact details for someone who could be contacted in the event of distress should also be given at the end of the questionnaire.
- The Consent Form should include the following standard paragraph: 'I understand that my medical notes and data collected during the study may be looked at by individuals from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to the records'.

It is responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

You should notify the REC in writing once all conditions have been met (except for site approvals from host organisations) and provide copies of any revised documentation with updated version numbers. Confirmation should also be provided to host organisations together with relevant documentation

Approved documents

The documents reviewed and approved at the meeting were:

Document	Version	Date
Covering Letter from Edward Roddy		29 March 2012
REC application: 89315/309525/1/664		02 April 2012
Protocol	1.0	15 February 2012
Investigator CV Dr Edward Roddy		29 March 2012
Investigator CV Dr Priyanka Chandtratre		29 March 2012
Investigator CV Professor Christian Mallen		05 April 2012
Investigator CV Jane Crompton Richardson		
Letter from Sponsor from Rhian Hughes		02 March 2012
Letter from Statistician from Dr Sara Muller		16 February 2012
Peer Review Report- Outcome letter and revisions made after initial review		24 February 2012
Letter to REC from Peer Review Committee from Professor PMS O'Brien		24 February 2012
GP Information Leaflet	1.0	15 February 2012
Letter of invitation to participant: Focus group interview	1.0	15 February 2012

appointment letter		
Letter of invitation to participant: invitation letter sent regarding the focus group interview		
Participant Information Sheet: for gout interview	1.0	15 February 2012
Participant Information Sheet: for the gout study	1.0	15 February 2012
Participant Consent Form: focus group interview	1.0	15 February 2012
Two week reminder letter for focus group interviews	1.0	15 February 2012
Covering letter sent with baseline questionnaire booklet	1.0	15 February 2012
Covering letter to be sent with 6 months follow up questionnaire booklet	1.0	15 February 2012
Covering letter to be sent with 12 months follow up questionnaire booklet	1.0	15 February 2012
Covering letter to be sent with 24 months follow up questionnaire booklet	1.0	15 February 2012
Covering letter to be sent with 36 months follow up questionnaire booklet	1.0	15 February 2012
Two week reminder postcard	1.0	15 February 2012
Two week reminder postcard sent after baseline questionnaire	1.0	15 February 2012
Reminder letter sent with repeat Questionnaire at week 4	1.0	15 February 2012
Reminder letter to be sent with repeat 6 months Questionnaire	1.0	15 February 2012
Reminder letter to be sent with repeat 12 months Questionnaire	1.0	15 February 2012
Reminder letter to be sent with repeat 2 years Questionnaire	1.0	15 February 2012
Reminder letter to be sent with repeat 3 years Questionnaire	1.0	15 February 2012
Focus Group reminder postcard	1.0	15 February 2012
Gout Study Questionnaire- baseline questionnaire booklet	1.0	15 February 2012
Gout Study Questionnaire- 6 months follow up questionnaire booklet	1.0	15 February 2012
Gout Study Questionnaire- 12 months follow up questionnaire booklet	1.0	15 February 2012
Gout Study Questionnaire- 2 years follow up questionnaire booklet	1.0	15 February 2012
Gout Study Questionnaire- 3 years follow up questionnaire booklet	1.0	15 February 2012
Evidence of insurance or indemnity from George Smith		21 July 2011

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

NHS
Health Research Authority

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

Further information is available at National Research Ethics Service website > After Review

12/NW/0297	**Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project

Yours sincerely

H Penistone

On behalf of
Mrs Jean Harkin
Chair

Email: helen.penistone@northwest.nhs.uk

Enclosures: List of names and professions of members who were present at the meeting and those who submitted written comments

"After ethical review – guidance for researchers"

Appendix 9: Summary table of questionnaire content (concept, operationalisation of concept, items and time-points)

Adapted from Chandratre et al (2012)

Gout

Concept	Operationalisation of concept	Empirical measure	Time-point
Frequency of gout flares	Number of flares in previous 12 months at baseline, 24 and 36 months, in previous 6 months at 6 and 12 months	5-point ordinal scale (0 to ≥5)	All
Age at gout diagnosis	Age in years	Numerical free text box	Baseline
Occurrence of current gout flare	Acute gout flare at time of completion of questionnaire	Yes/No	All
History of oligo/ polyarticular flares	Reported experience of gout in more than one joint	Yes/No	All
Use of allopurinol	Reported use of allopurinol	Yes/No	All
	Reported dose of allopurinol	10 categories (9 dose categories, 1 option if dose unknown). Free text box for dose.	All

How gout affects life

Concept	Operationalisation of concept	Empirical measure	Time-point
Disease-specific HRQOL	Gout Impact Scale (GIS) (Hirsch et al 2008) GIS subscales:	24 items Response via 5-point ordinal scale	All
Gout concern overall, Gout medication side effects, Unmet gout treatment needs, Wellbeing during an attack, Concern during an attack	GIS Concern overall (GIS CO), GIS Medication side effects (GIS MSE), GIS Unmet treatment needs (GIS UTN), GIS Wellbeing during an attack (GIS WBDA), GIS Concern during an attack (GIS CDA)		
Illness perception	Questions from Revised Illness Perception Questionnaire (IPQ-R) (Moss-Morris et al 2002)	4 items Response via 5- point Likert scale	All

General health

Concept	Operationalisation of concept	Empirical measure	Time-point
Physical function	36-Item Short-Form Health Survey Physical Function Subscale (SF-36 PF 10) (Ware & Sherbourne, 1992)	10 items Response via 3-point ordinal scale	All
Activity limitation/disability	Health Assessment Disability Index (HAQ-DI) (Bruce & Fries, 2003)	20 items Response via 4-point ordinal scale Additional question regarding the use of aids, devices or assistance	All
Pain severity	Numerical rating scale (NRS) Pain (Bruce & Fries, 2003)	11-point NRS	All
Global assessment of health	Numerical rating scale (NRS) Patient Global (Bruce & Fries, 2003)	11-point NRS	All
Comorbidities	Diabetes, Hypertension (HT), Hyperlipidaemia (HL), Myocardial infarction (MI), Angina, Cerebrovascular accident (CVA), Transient Ischaemic attack (TIA), Renal failure (RF), Renal calculi	Yes/No	Baseline only

How participants feel

Concept	Operationalisation of concept	Empirical measure	Time point
Depression	Patient Health Questionnaire-9 (PHQ-9) (Kroenke, Spitzer & Williams, 2001)	9 items Response via 4-point ordinal scale Additional question if subject check any of 9 items above. Response via 4-point ordinal scale	Baseline 12 months 36 months
Anxiety	Generalised Anxiety Disorder-7 (GAD-7) (Spitzer et al 2006)	7 items Response via 4-point ordinal scale	Baseline 12 months 36 months

Foot and other joint problems

Concept	Operationalisation of concept	Empirical measure	Time point
Hallux Valgus	Line drawings of left and right feet. (Roddy, Zhang & Doherty, 2007d)	5 line drawings of left and right feet displaying increase in severity of HV	Baseline 12 months 36 months
Foot pain, aching & stiffness	Frequency of foot pain or aching or stiffness. (Dufour et al 2009)	1 item Response via 5-point ordinal scale	Baseline 12 months 36 months
Foot pain location	Location of foot pain on manikin illustration (Garrow, Silman & Macfarlane, 2004)	Foot manikin illustrations	Baseline 12 months 36 months
Foot function	Manchester Foot Pain and Disability Index (Garrow et al 2000)	17 Items Response via 3-point Likert scale	Baseline 12 months 36 months
Health care consultations regarding feet	Consultant with GP relating to feet NHS or Private consultation with Physiotherapist, Podiatry & Chiropody relating to feet	1 Item Yes/No 6 nominal response categories	Baseline 12 months 36 months
Pain	Pain, ache, discomfort or stiffness lasting longer than one day in any part of their body during the past 4 weeks Location of pain on a body manikin (Hunt et al 1999; Lacey et al 2005)	1 Item Yes/No Body manikin	Baseline 12 months 36 months Baseline 12 months 36 months

How gout affects work

Concept	Operationalisation of concept	Empirical measure	Time-point
Occupational characteristics	Current employment status	11 nominal categories	Baseline 12 months 36 months
	Taken time off work due to gout in last 6 months	Yes/No	Baseline 12 months 36 months
	Current abilities to undertaken usual employment	5 nominal categories	Baseline 12 months 36 months
	Is inability to undertake normal employment attributed to joint problems	Yes/No	Baseline 12 months 36 months

Demographics and socioeconomics, BMI and alcohol

Concept	Operationalisation of concept	Empirical measure	Time-point
Date of birth	Date of birth	Date of birth	Baseline
Sex	Sex	Male/ Female	Baseline
Relationship status	Relationship status	6 nominal categories	Baseline
Education	Attendance at Higher Education	Yes/No	Baseline
Ethnic origin	Ethnic origin	6 nominal categories	Baseline
Anthropometry	Height Weight	Metres/Feet & Inches Kilograms/Stones & pounds	All All
Alcohol consumption	Frequency of alcohol consumption Quantity of alcohol of specified type and volume consumed	7 ordinal categories Free text	Baseline Baseline

Appendix 10: Descriptions of model fit indices included in thesis

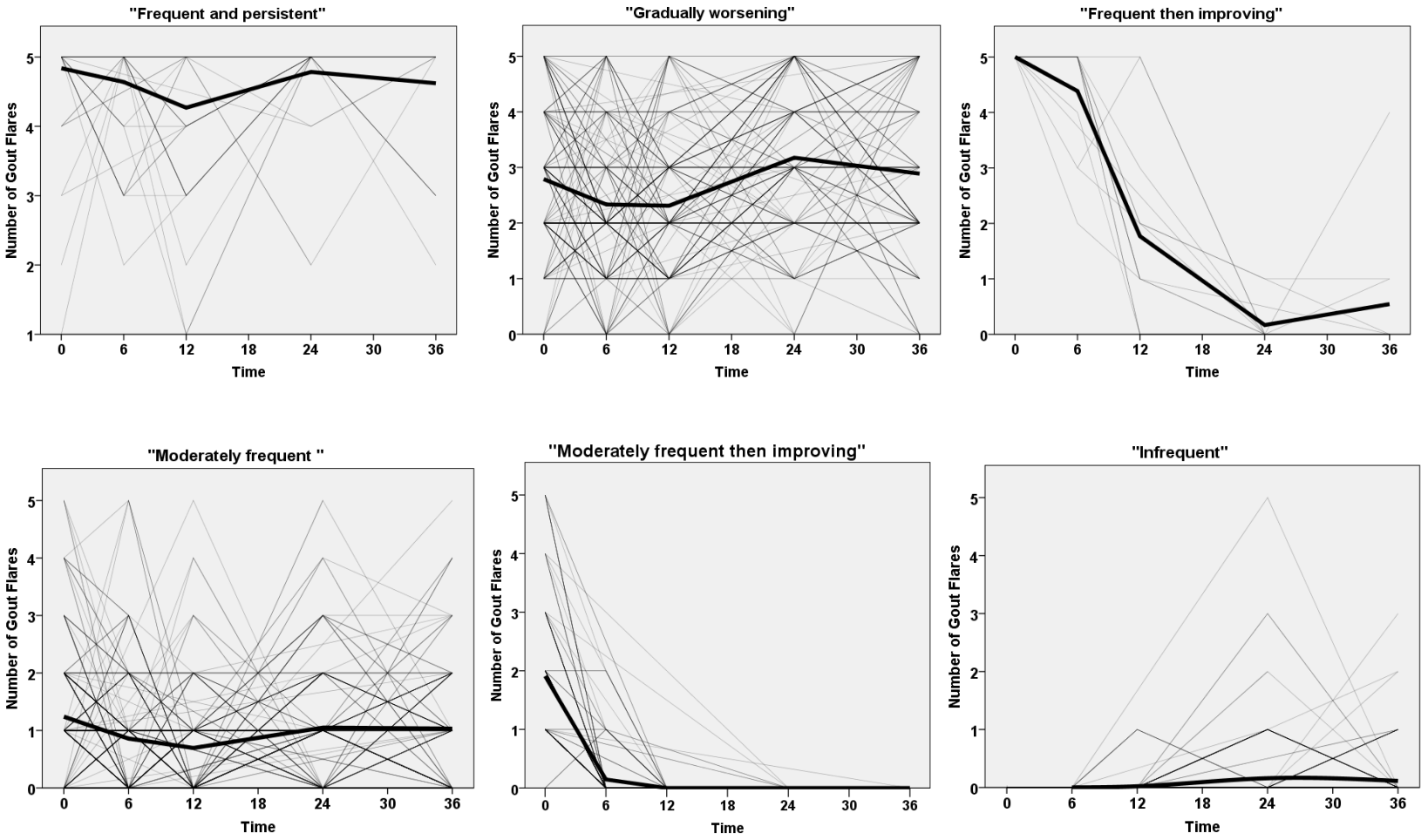
Chapter 5 Latent growth curve modelling

Indices	Description/comments
Chi-Squared Statistic (χ^2)	The **Chi-Squared Statistic** assesses the magnitude of discrepancy between the covariance matrices of the observed sample and fitted model (Hu & Bentler, 1999). The chi-squared statistic is sensitive to sample size (Hooper, Coughlan & Mullen, 2008; Wickrama et al 2016) and where the sample size is large a model will nearly always be rejected (Bentler & Bonnett, 1980).
Comparative Fit Index (CFI) (Bentler, 1990)	The **Comparative Fit Index** is an incremental fit index which evaluates the fit of the model compared to reduced nested null model which assumes that all the latent variables are uncorrelated (Hooper, Coughlan & Mullen, 2008; Hu & Bentler, 1998). The CFI takes account of sample size and is thus less effected by sample size than χ^2 (Hooper, Coughlan & Mullen, 2008).
Tucker-Lewis Index (TLI) (Tucker & Lewis, 1973)	The **Tucker-Lewis index** is an incremental fit index which also evaluates the fit of the model compared to a nested model which assumes that all the latent variables are uncorrelated (Hooper, Coughlan & Mullen, 2008; Wickrama et al 2016). The TLI also considers the complexity of the model by consideration of the number of degrees of freedom (Kenny, 2015; Wickrama et al 2016).
Root mean square error of approximation (RMSEA) (Steiger & Lind, 1980)	The **Root mean square error of approximation** is an absolute fit indices which evaluates how well the model fits the sample data whilst taking account of model complexity (Hooper, Coughlan & Mullen, 2008; Wickrama et al 2016). As a model fit index the RMSEA is fairly insensitive to sample size (Wickrama et al 2016). The result obtained from the RMSEA is accompanied with a confidence interval which can aid interpretation regarding the adequacy of model fit (Hooper, Coughlan & Mullen, 2008).
Weighted root-mean-square residual (WRMR)	The **Weighted root-mean-square residual** measures the weighted average differences between the variances and covariances of the sample and the fitted model (Yu, 2002). The appropriate cut off for the WRMR is increased for more time-points and increased sample size (Yu, 2002).

Criterion	Description/comments
BIC	**Bayesian information criteria** (BIC) is based on the log likelihood for a model but then penalises for the number of free (estimated) parameters in the model and also the sample size (Feldman, Masyn & Conger, 2009; Nylund, Asparouhov & Muthen, 2007). BIC is the most commonly used information criterion when determining the optimal number of latent classes (Feldman, Masyn & Conger 2009; Twisk & Hoekstra, 2012).
AIC	The **Akaike information criteria** (AIC) (Akaike, 1987) is an information criterion which based on the log likelihood for the model but then penalises for the number of free (estimated) parameters in the model (Nylund, Asparouhov & Muthen, 2007).
LMR-LRT	Likelihood ratio tests (LRTs) compare a current model (*k*-class) with a model which has one less class than the current model (*k*-1) (Feldman, Masyn & Conger, 2009; Jung & Wickrama, 2007; Nylund, Asparouhov & Muthen, 2007; Wickrama et al 2016). The **Lo-Mendell –Rubin Likelihood Ratio Test** (LMR-LRT) was developed by Lo, Mendell & Rubin (2001) and uses an approximate sampling distribution for the likelihood ratio test (Peugh & Fan, 2012). The LMR-LRT compares the fit of adjacent class models, i.e. models which have *K*-1 and *K* classes, and generates a p value which indicates whether there is a statistically significant improvement in model fit of the *K* class model versus the previous *K*-1 class mode as displayed in table (Nylund, Asparouhov & Muthen, 2007; Wickrama et al 2016).
BLRT	The **Bootstrapped Likelihood Ratio Test (BLRT)** was developed by McLachlan & Peel (2000) to estimate the distribution of the log likelihood ratio test utilising bootstrap samples. BLRT compares the fit of adjacent class models, i.e. models which have *K*-1 and *K* classes, and generates a p value which indicates whether there is a statistically significant improvement in model fit of the K class model versus the previous *K*-1 class model (Nylund, Asparouhov & Muthen, 2007; Wickrama et al 2016).
Average posterior probabilities	**Posterior probability** is calculated from the observed variable using maximum likelihood estimation and it is the probability of an individual belonging to any of the (latent) classes (Wickrama et al 2016). As part of the LCGA individuals are assigned to the latent class which is associated with the highest posterior probability (Berlin, Parra & Williams, 2104; Wickrama et al 2016). The average posterior probability of **each latent** class is calculated in LCGA (Muthen & Muthen, 2000; Wickrama et al 2016).
Entropy	**Entropy** is a standardised index of the classification accuracy of a model, with values for entropy ranging from 0 to 1 (Peugh & Fan 2012; Wickrama et al 2016). Calculation of entropy involves information about the posterior probabilities of individuals within classes and provides an **overall entropy result for the latent class model** (Clark & Muthen, 2009; Feldman, Masyn & Conger, 2009).

Indices	Description/comments
Intraclass correlation coefficient (ICC)	The **intraclass correlation coefficient** is the ratio of the between subject variance to the total variance (which is a combination of the between and within subject variance). An ICC is equal to 1 when all the variance is explained by the differences between individuals (Kirwood & Sterne, 2003; West, Welch & Galecki, 2015).
Likelihood Ratio Test (LRT)	The **likelihood ratio test** compares the log likelihood from one model with the log likelihood from another model e.g. comparing a nested and reference model (Kirwood & Sterne, 2003; West, Welch & Galecki, 2015).
Wald χ^2 test	The **Wald χ^2 test** statistic is equal to the estimated regression coefficient divided by its standard error, and its square approximates the Chi-squared distribution (Petrie & Sabin, 2009). The Wald χ^2 test statistic is used to test the null hypothesis that a regression coefficient is zero (Petrie & Sabin, 2009; West, Welch & Galecki, 2015).
Akaike Information Criteria (AIC)	The **Akaike Information Criteria** (AIC) (Akaike, 1987) is an information criterion which is based on the log likelihood for the model but then penalises for the number of free (estimated) parameters in the model (Nylund, Asparouhov & Muthen, 2007).

Appendix 11: Spaghetti plots displaying the individual gout flares reported for each participant in each latent class with an interpolation line fitted

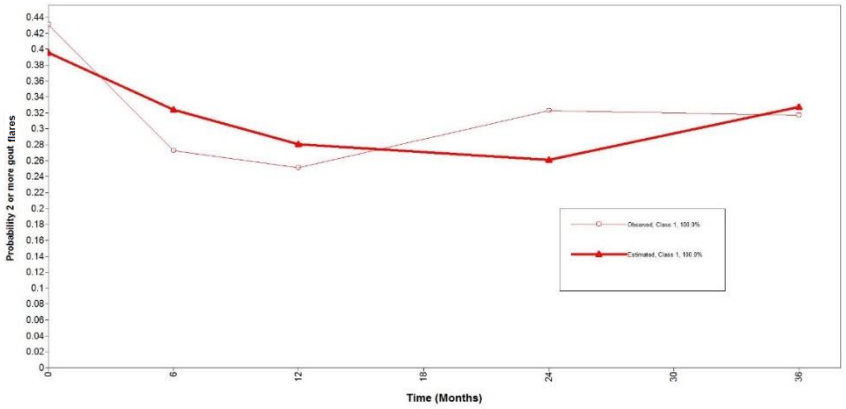


NB. 5 = ≥5 flares

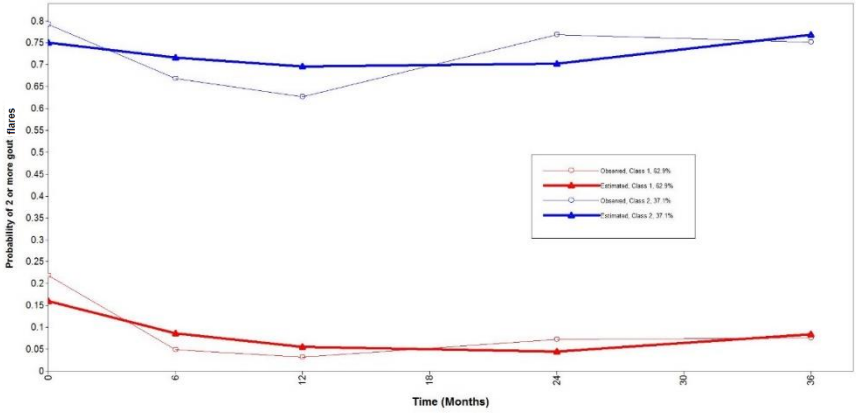
Appendix 12: Sensitivity analyses for quadratic ordinal LCGA models plots probability of ≥ 2 gout flares at each time-point

<u>For participants who responded at >3 time-points)</u>

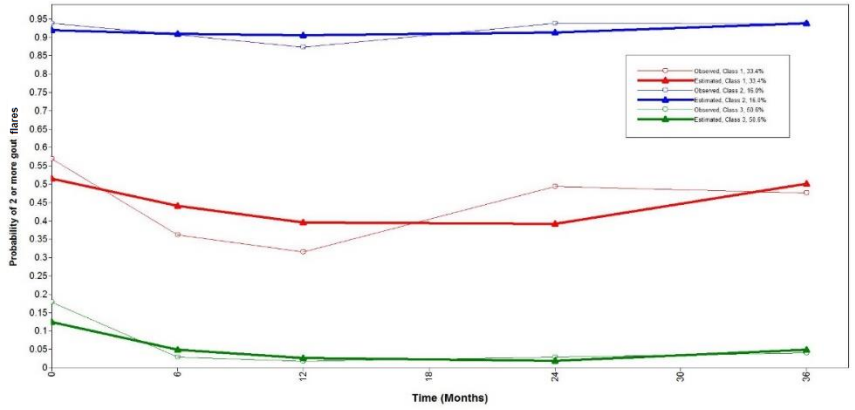
One class solution



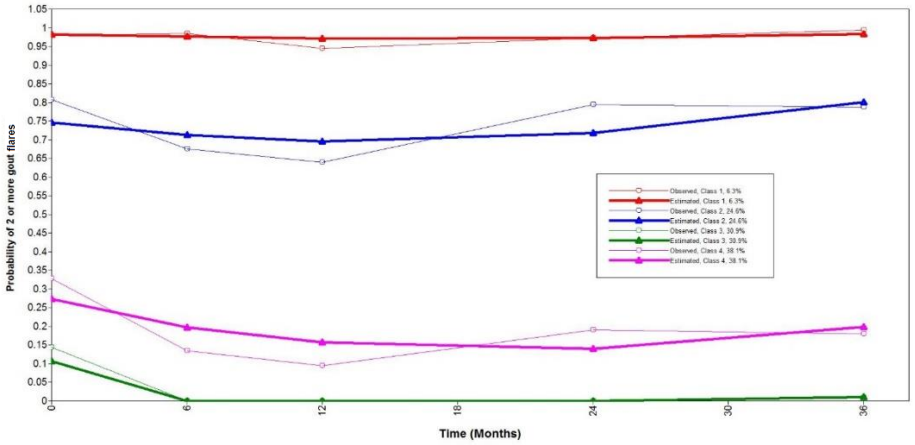
Two class solution



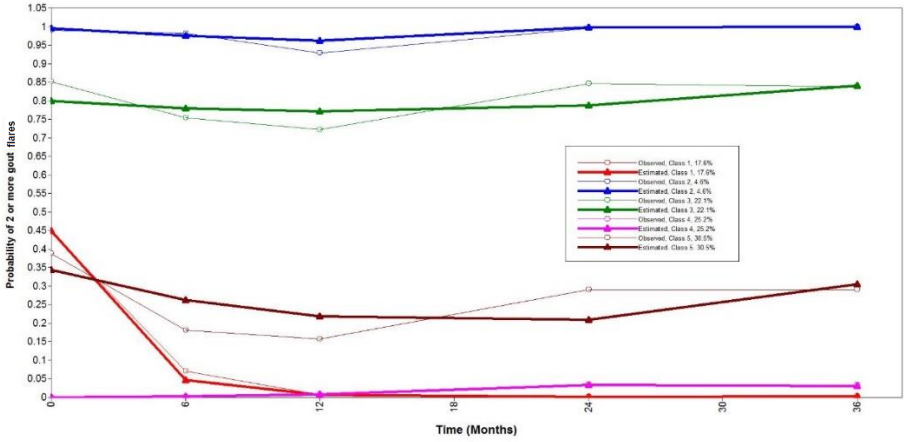
Three class solution



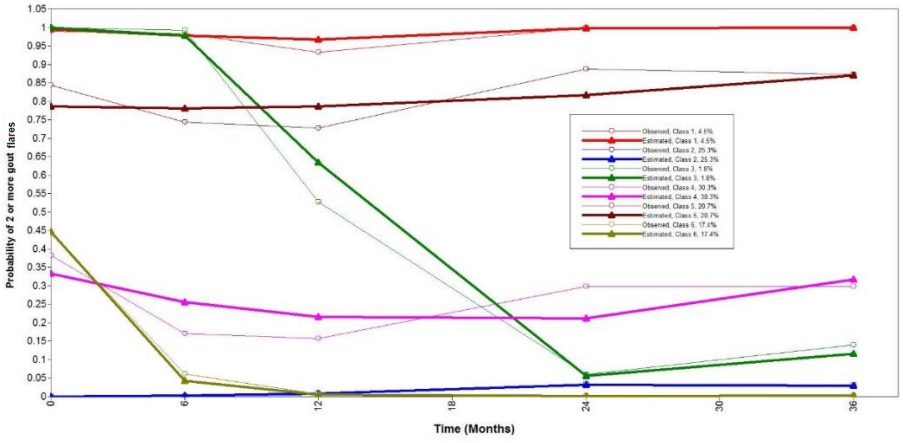
Four class solution



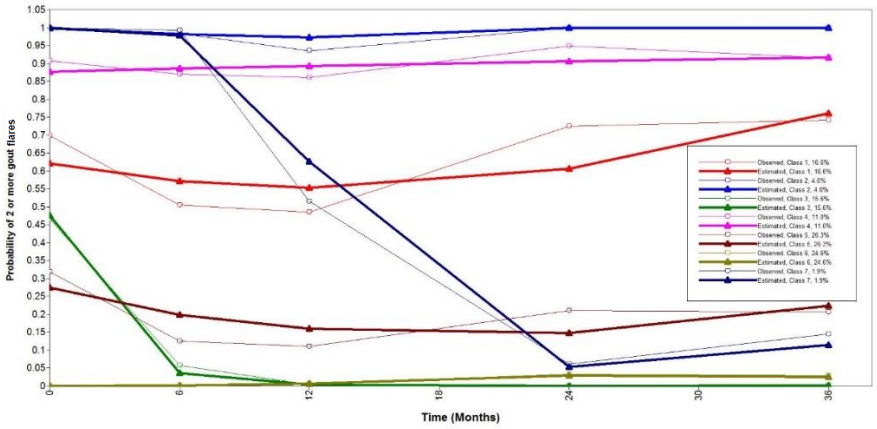
Five class solution



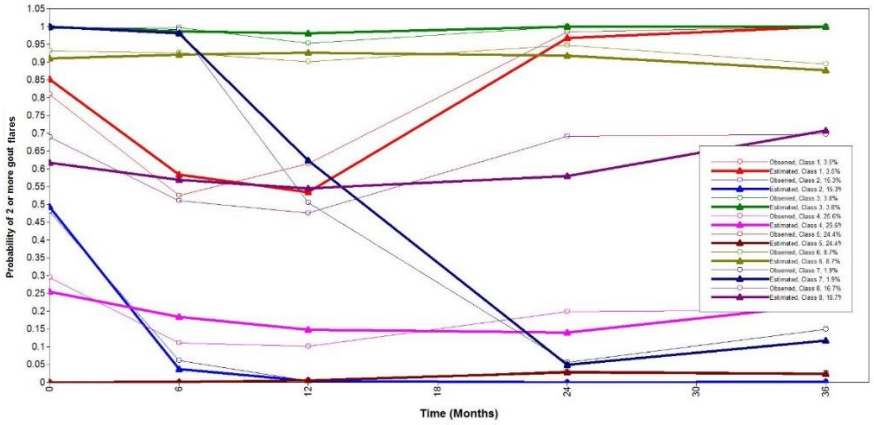
Six class solution



Seven class solution



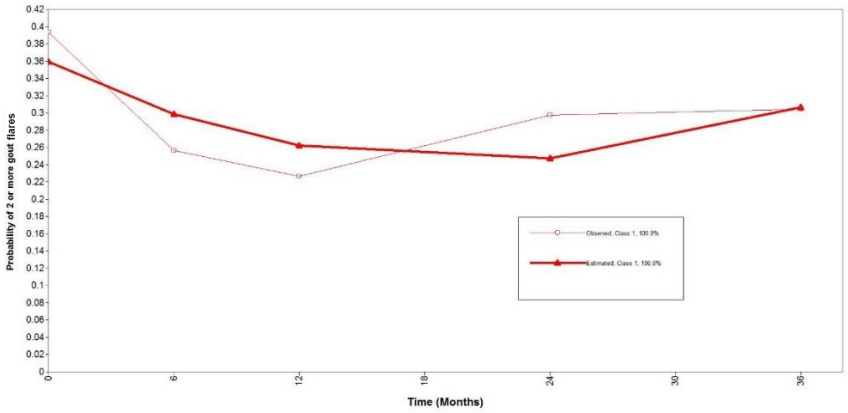
Eight class solution



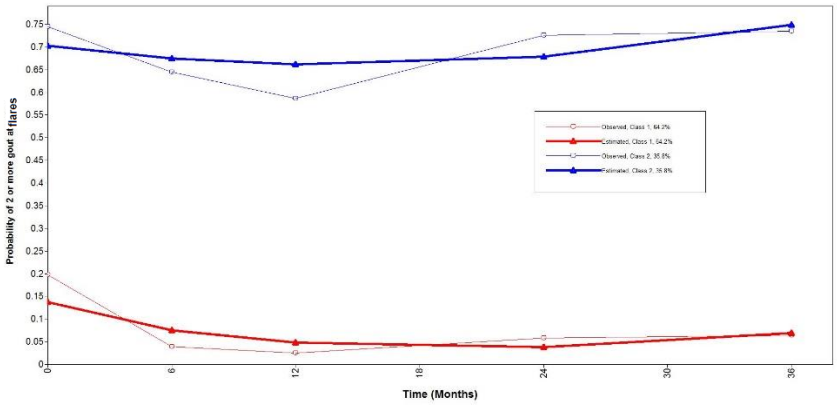
Appendix 12 cont. Quadratic ordinal LCGA model plots; probability of ≥2 gout flares at each time-point

For participants who responded at 5 time-points

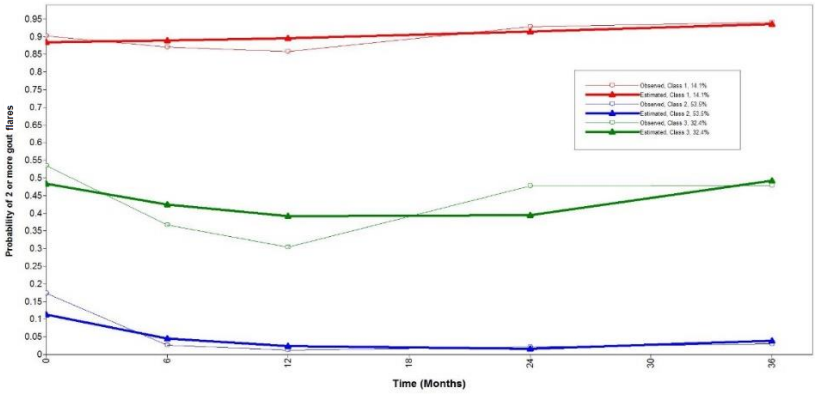
One class solution



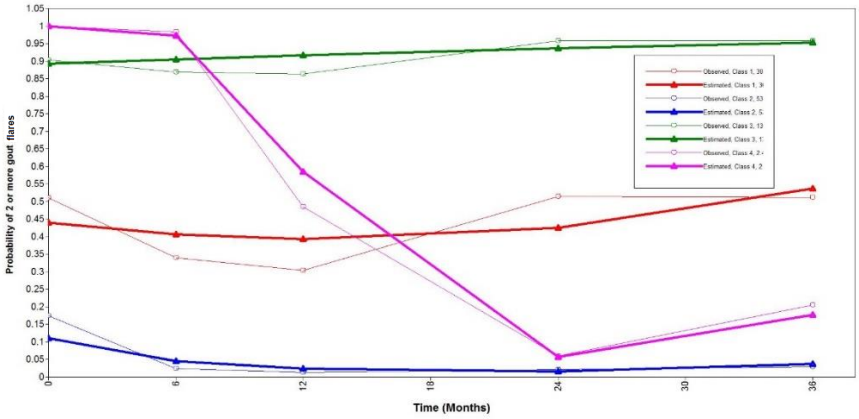
Two class solution



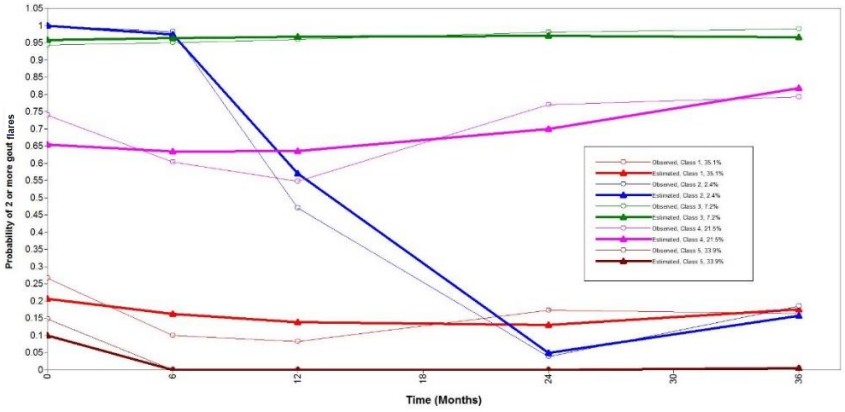
Three class solution



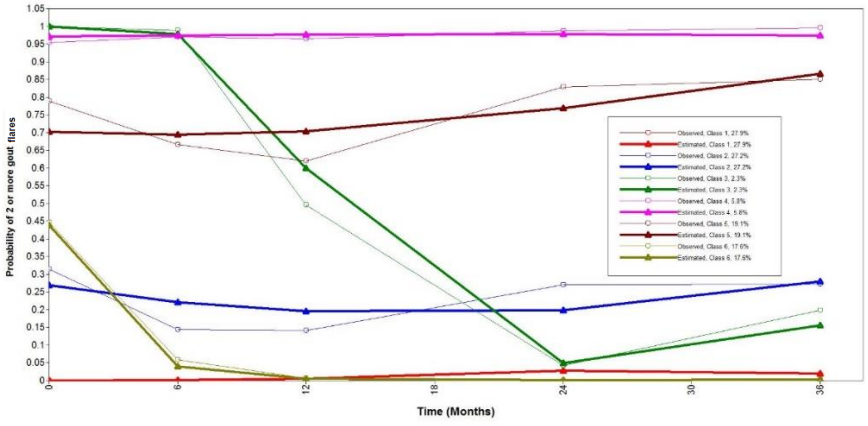
Four class solution



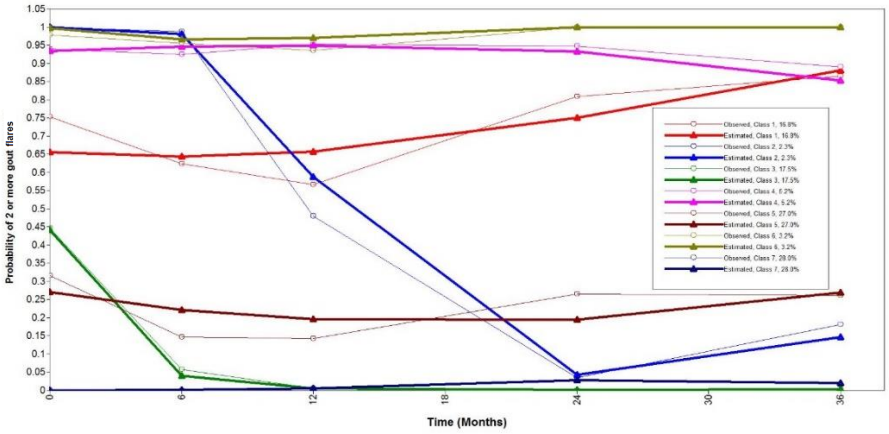
Five class solution



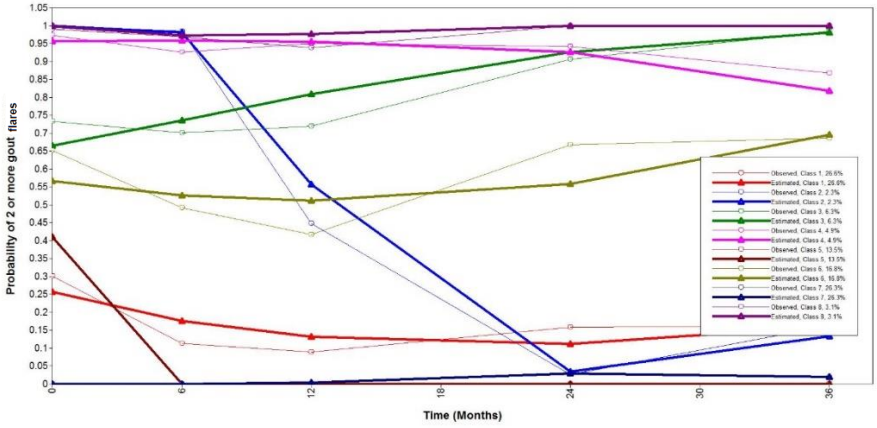
Six class solution



Seven class solution



Eight class solution



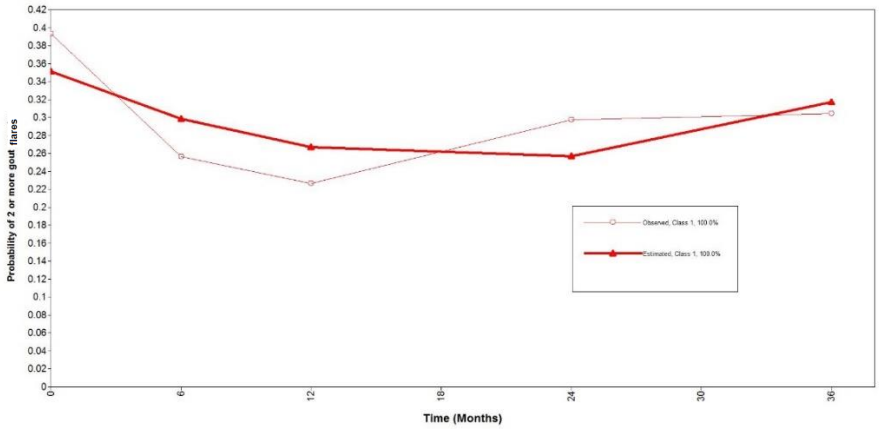
Appendix 13: Model fit indices for GMM and the plots for the GMM dataset consisting of participants who self-reported gout flare at five time-points

Number of classes	AIC	BIC	BLRT	LMR-LRT	Entropy	Average posterior probabilities	Number per class (%) based on most likely class membership
1	5127.225	5159.864	-	-	-	-	437 (100)
2	5021.396	5070.356	P value <0.001	P value <0.001	0.610	0.846/0.873	80/357 (18.3/81.7)
3[b]	5017.517	5082.796	P value 0.0515[b]	P value 0.1889	0.741	0.676/0.979/0.830	178/248/11 (40.7/56.8/2.5)
4[b,c]	5013.749	5095.348	P value 0.0538[b]	P value 0.1223	0.526	0.889/0.761/0,735/0.589	10/120/131/176 (2.3/27.5/30/40.2)
5[b]	5012.178	5110.097	P value 1.000[b]	P value 0.1550	0.581	0.640/0.872/0.691/0.766/707	175/11/16/133/102 (40.1/2.5/3.7/30.4/23.3)
6	-	-	-	-	-	-	-
7	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-

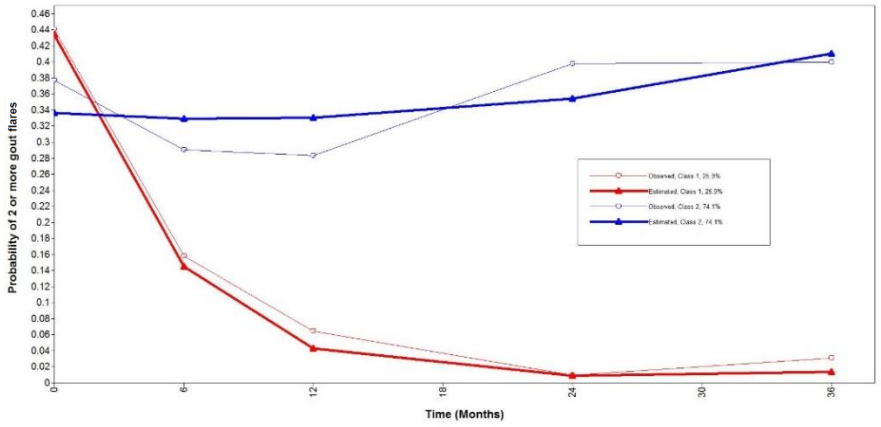
AIC, Akaike Information Criteria; **BIC**, Bayesian information criteria; **BLRT**, Bootstrap likelihood ratio test; **LMR-LRT**, Lo Mendell Rubin likelihood ratio test.

[a] model fixed to avoid singularity, warning given that model may not be identified, [b] Warning that not all of bootstraps converged, [C] Warning that some draws had smaller LRT than observed LRT **Table formatted as per Guidelines for Reporting on Latent Trajectory Studies (GRoLTs)** (van de Schoot et al 2017)

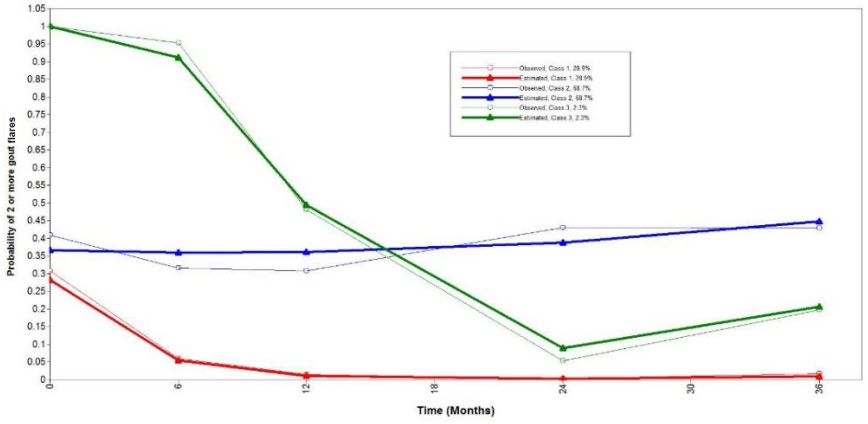
One class solution GMM



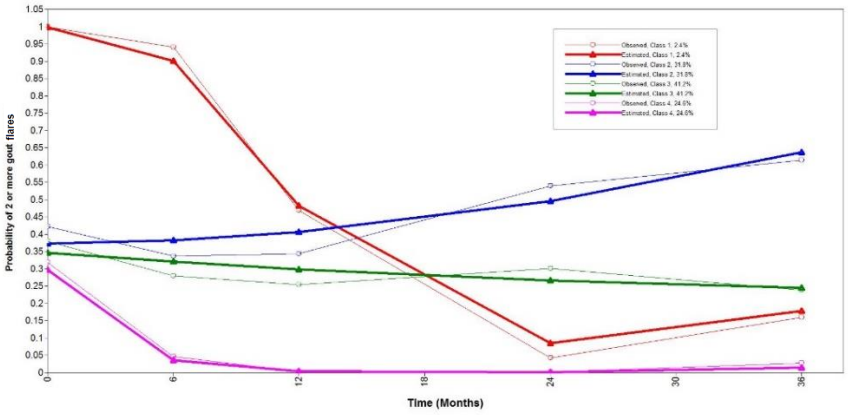
Two class solution GMM



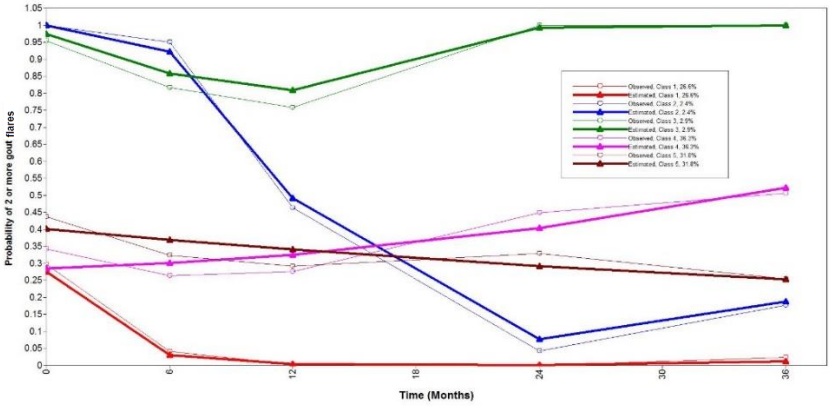
Three class solution GMM



Four class solution GMM

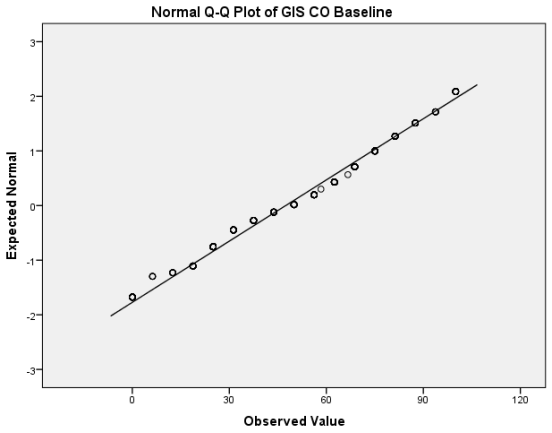


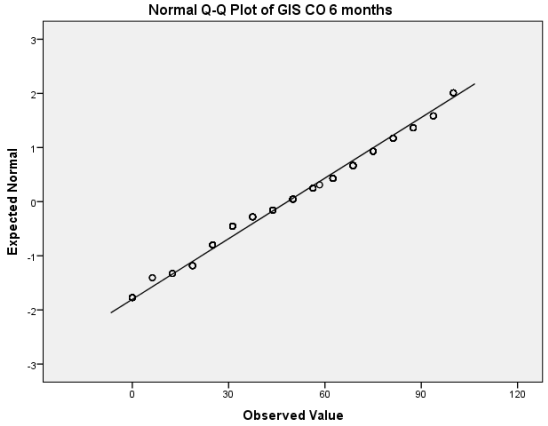
Five class solution GMM

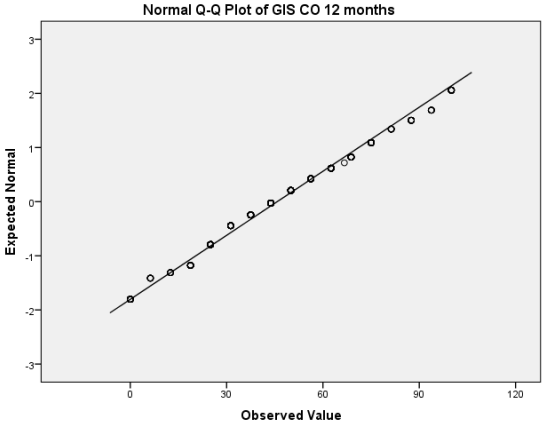


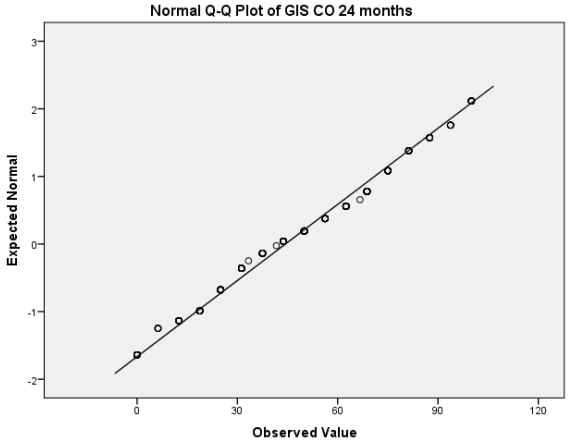
Appendix 14: Q-Q plots of all five HRQOL outcomes at all five time-points

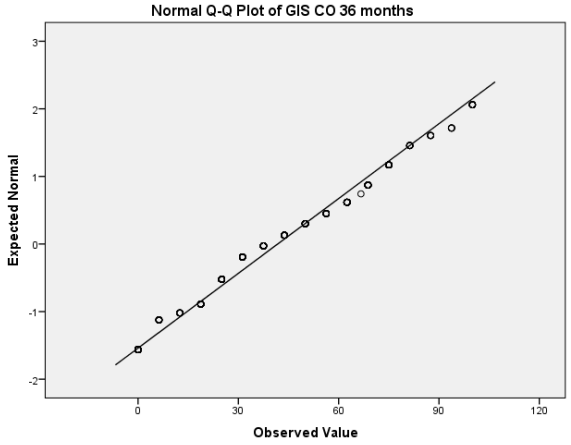
GIS CO Q-Q plots at each time-point (baseline, 6, 12, 24 and 36 months)



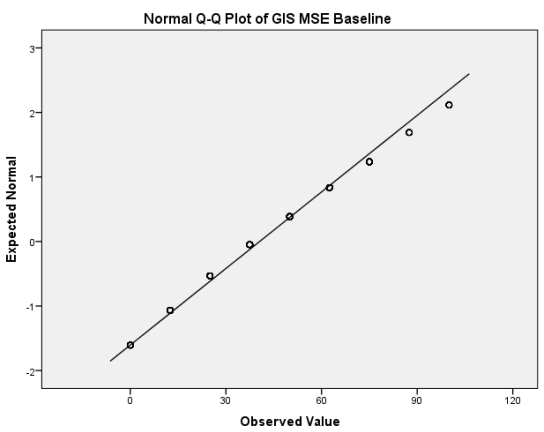


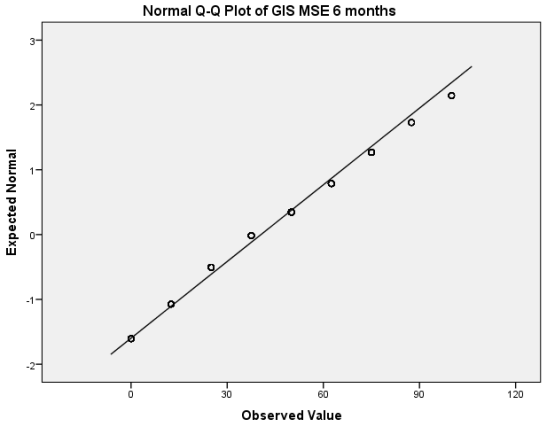


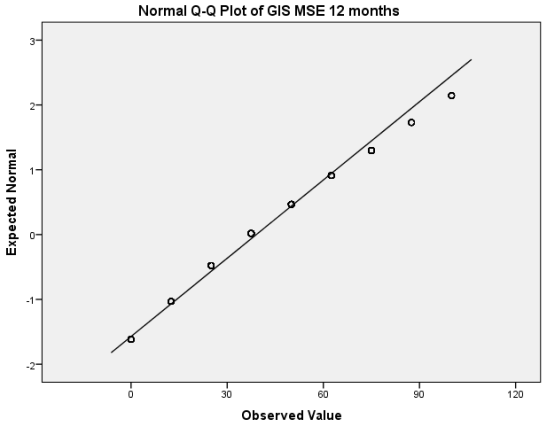


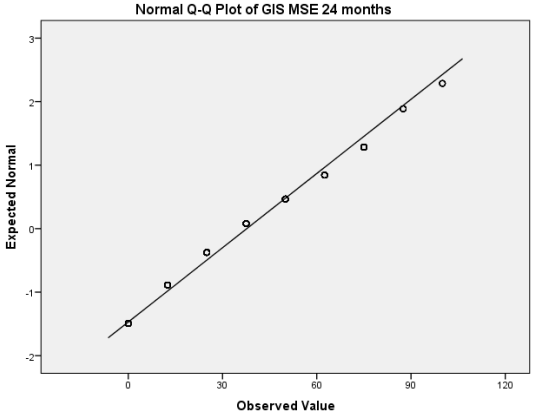


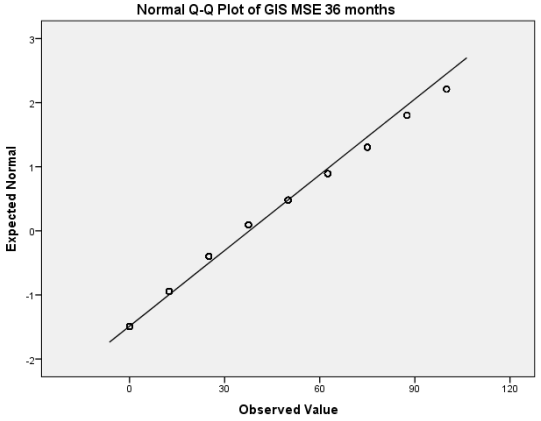
GIS MSE Q-Q plots at each time-point (baseline, 6, 12, 24 and 36 months)



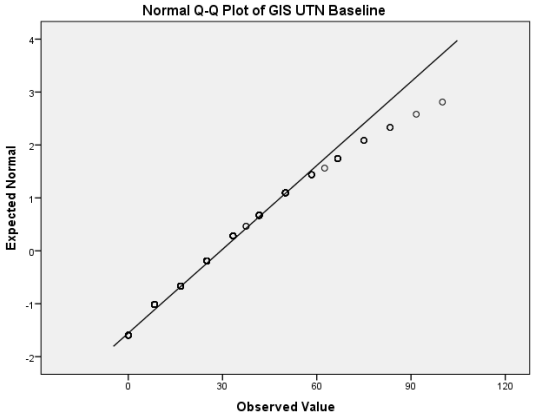


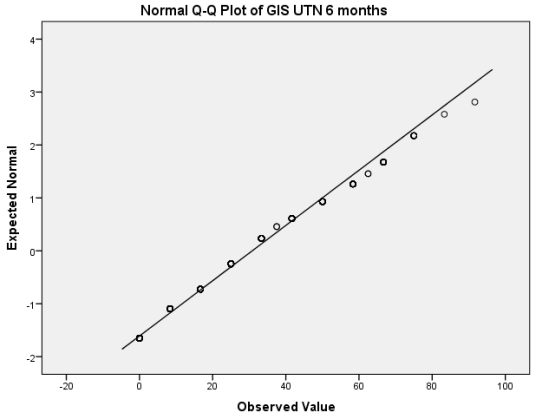


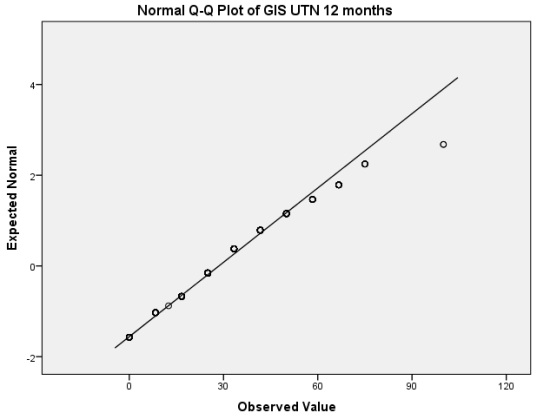


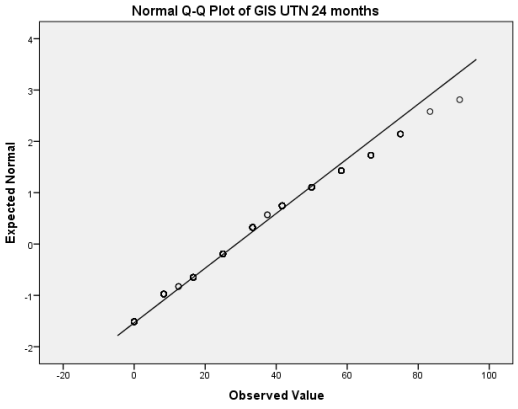


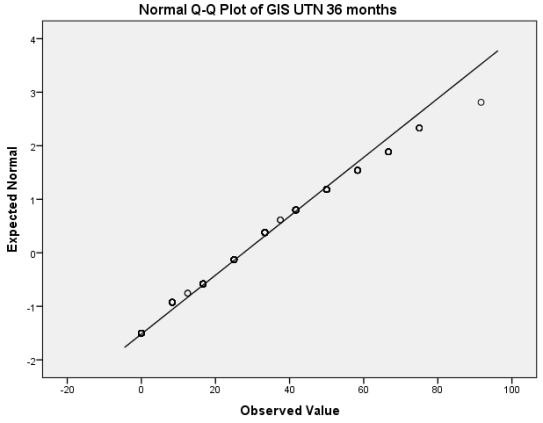
GIS UTN Q-Q plots at each time-point (baseline, 6, 12, 24 and 36 months)



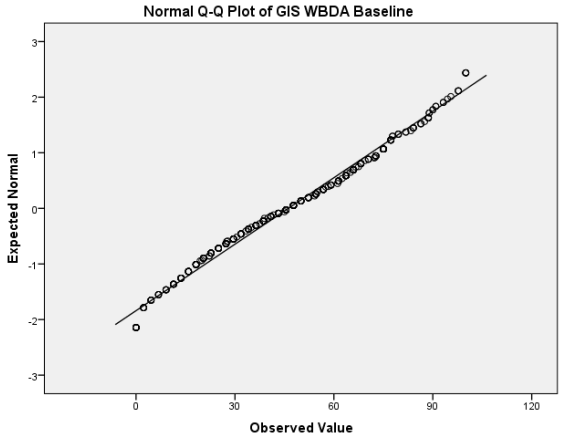


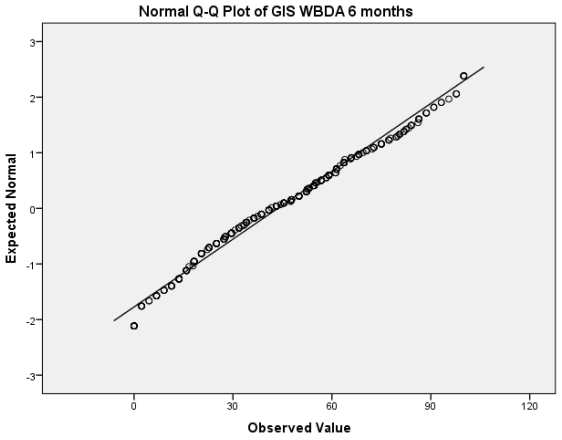


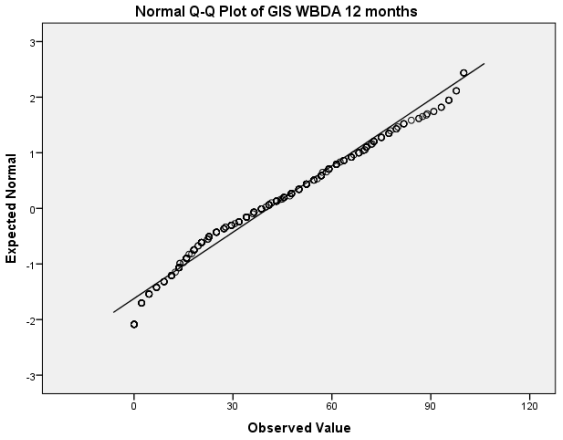


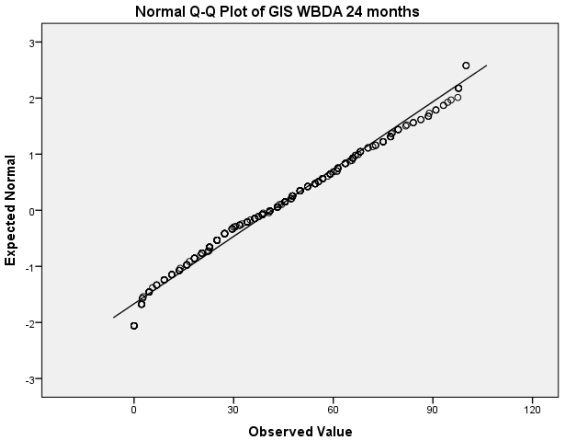


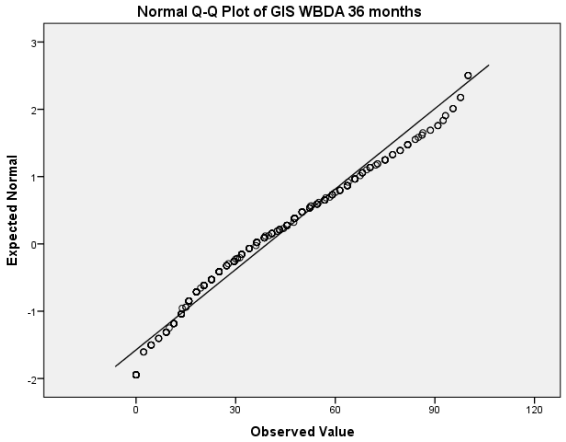
GIS WBDA Q-Q plots at each time-point (baseline, 6, 12, 24 and 36 months)



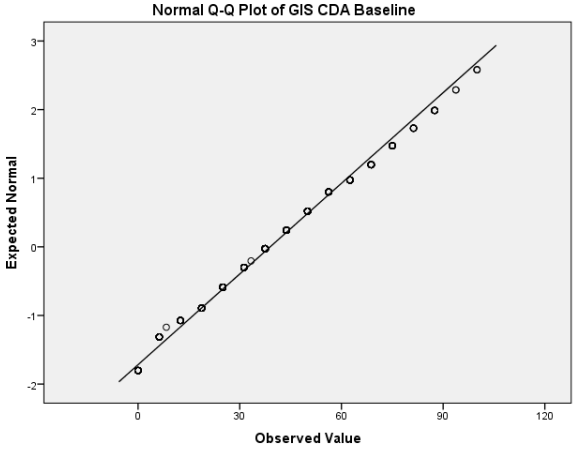


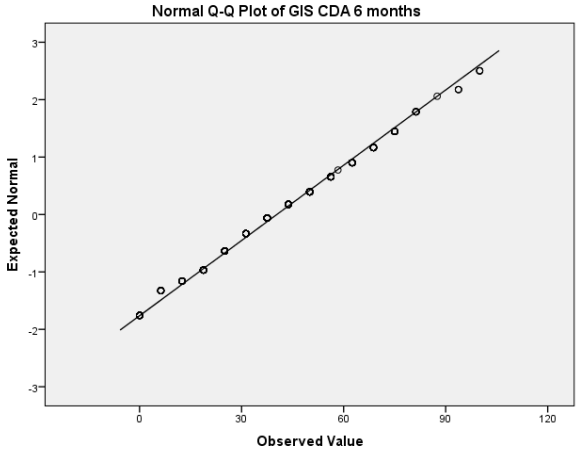


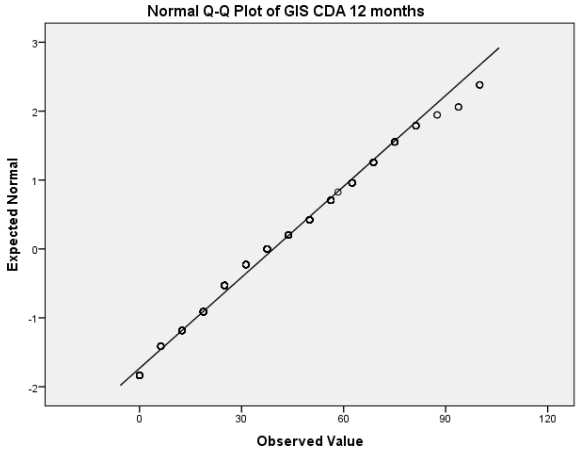


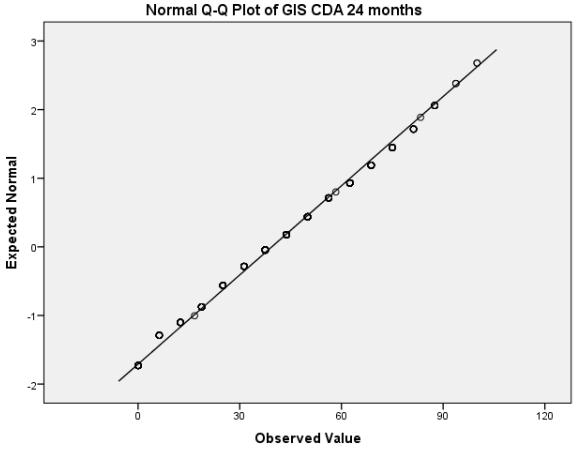


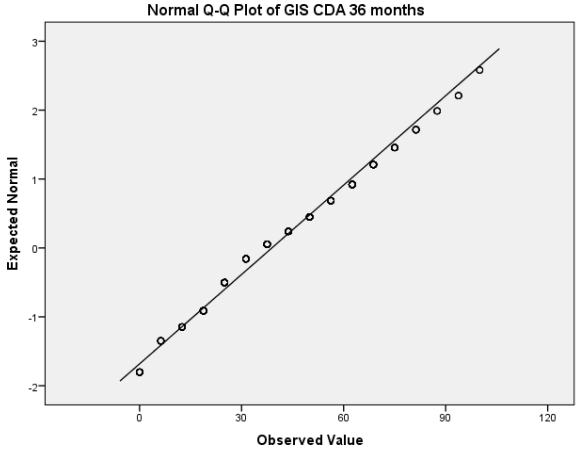
GIS CDA Q-Q plots at each time-point (baseline, 6, 12, 24 and 36 months)



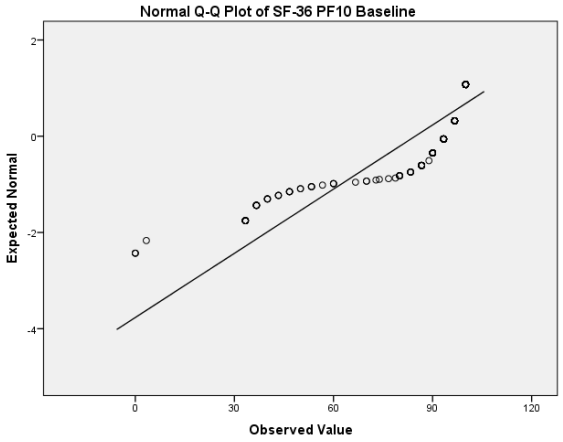


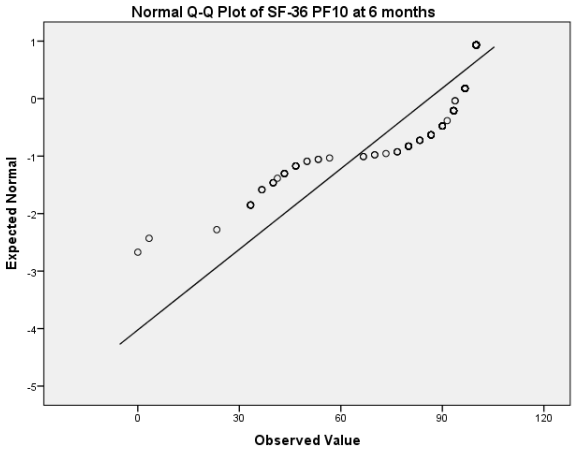


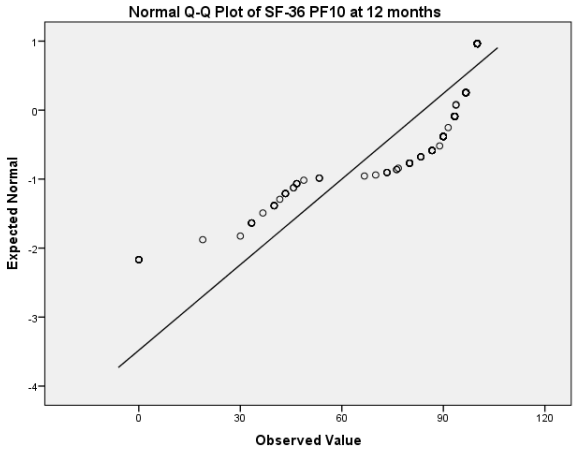


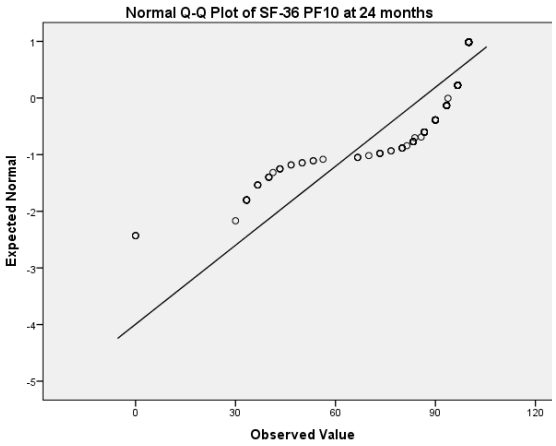


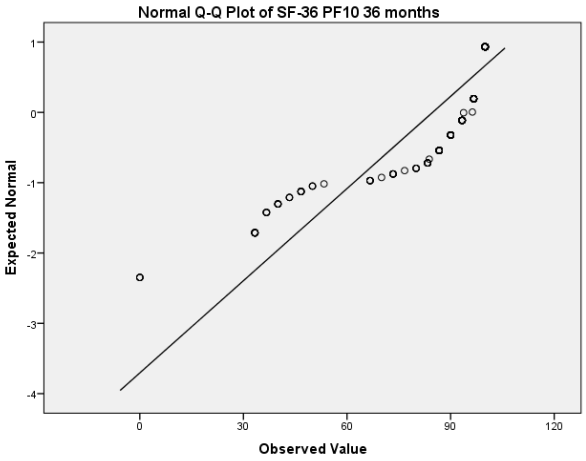
SF-36 PF10 Q-Q plots at each time-point (baseline, 6, 12, 24 and 36 months)



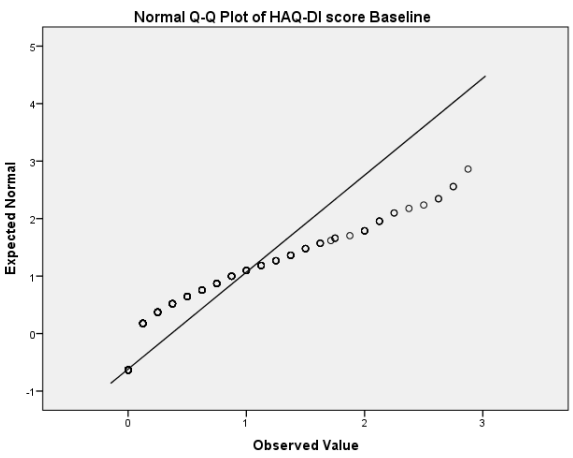


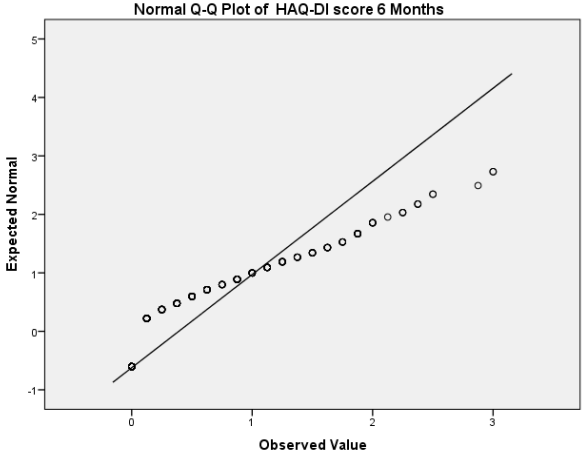


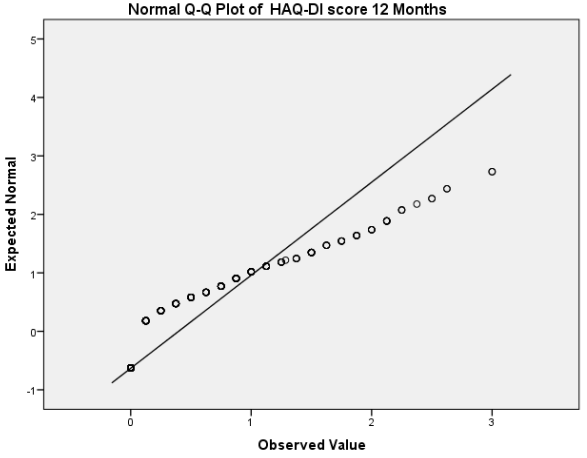


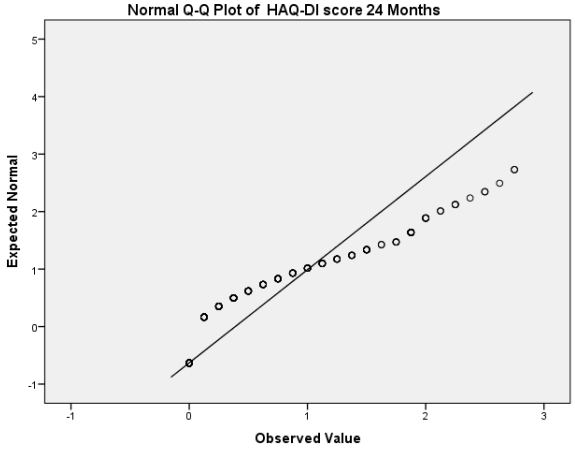


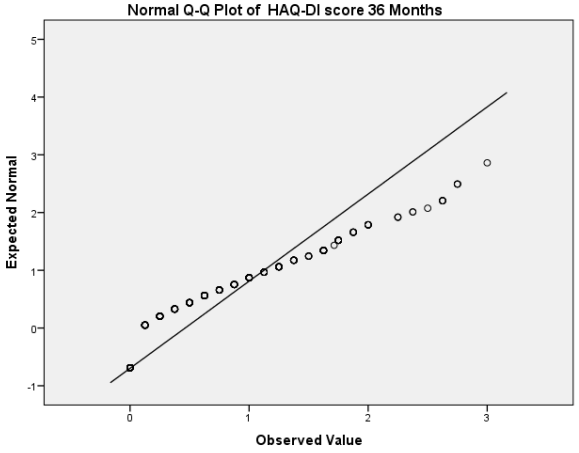
HAQ-DI Q-Q plots at each time-point (baseline, 6, 12, 24 and 36 months)











Appendix 15: Covariance structures trialled during linear mixed modelling

Structure	Description
independent	One unique variance parameter per random effect, all covariance 0; the default unless notation is used
exchangeable	Equal variances for random effects, and one common pairwise covariance
identity	Equal variances for random effects, all covariance 0
unstructured	All variances and covariances to be distinctly estimated
Toeplitz 1	Assumes that within-group errors have a Toeplitz structure for which correlations are constant with respect to time lags less than or equal to 1 and are 0 for lags greater than 1.
autoregressive 1	Assumes that within-group errors have an autoregressive structure of order 1.

p. 287 & 289

(StataCorp, 2013)