University of Nottingham

UK | CHINA | MALAYSIA

An investigation into inflammasome-gene expression profile and genome-wide genotype data in gout

Gabriela Sandoval-Plata, BSc, MSc

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

Academic Rheumatology

School of Medicine, University of Nottingham

Declaration

I herby declare that this thesis is the result of original research. It has been conducted by myself under the supervision of Professor A. Abhishek and Professor Kevin Morgan. This thesis has not been accepted elsewhere for any degree, diploma or other qualification. All the authors and works to which reference has been made, are fully acknowledged.

Abstract

Background: Gout is the most common form of inflammatory arthritis, with a prevalence of 2.49% in the United Kingdom. It is characterised by episodes of acute inflammation (flares) caused by shedding of monosodium urate (MSU) crystals that deposit within and around joints when serum urate (SU) increases above its solubilisation point. Despite the central role of hyperuricaemia in the pathogenesis of gout, not everyone with elevated SU develop symptomatic gout. The reasons underlying the transition from asymptomatic hyperuricaemia, with our without MSU crystal deposits, to symptomatic gout remain poorly understood.

Objectives: (1) To explore expression of inflammasome and toll-like receptors (TLRs)-associated genes and cytokine levels in different stages of the pathogenesis of gout: normal SU, hyperuricaemia without MSU crystal deposits, hyperuricaemia with asymptomatic MSU crystal deposits, intercritical gout and gout flares. (2) To generate a polygenic risk score (PRS) to distinguish gout cases from controls regardless of SU. (3) To generate and validate a genome-wide association study (GWAS) of asymptomatic hyperuricaemia controls vs. gout cases. (4) To generate a PRS to distinguish gout cases from controls with asymptomatic hyperuricaemia.

Methods: *Gene expression and cytokine profiling:* Participants recruited for three clinical studies conducted at the Department of Academic Rheumatology were classified according to their SU, presence of MSU crystals, and/or gout stage. A total of 108 were included in the gene expression analysis, which evaluated the relative expression of 86 genes associated to inflammasome and TLRs pathways, using the QIAGEN RT_2 RNA PCR Arrays. 185 participants

were included in the cytokine measurements, conducted using latex agglutination and Meso-Scale Discovery by Affinity Biomarkers Labs. mRNA and cytokine levels were compared among groups using the Kruskal-Wallis H Test with Bonferroni post-hoc for independent measurements, and Wilcoxon signed-rank rest or Friedman test with Bonferroni post-hoc for repeated measurements. P-values were corrected for multiple testing using a false discovery rate of 5%.

GWAS and PRS studies: Phenotype and genotype data from participants of the UK Biobank resource were used to derive two cohorts: (1) gout cases vs. non-gout controls (regardless SU levels), and (2) gout cases vs. asymptomatic hyperuricaemia controls. The second cohort was divided into the discovery dataset (comprised by 70% of the original cohort) and the replication dataset (comprised by 30% of the original cohort). Genotype data was quality controlled using PLINK v1.9 to remove participants with non-European ethnicity, third degree relatives, sex mismatches, low call-rate and heterozygosity outliers, and markers deviating from Hardy-Weinberg equilibrium and high missingness. Discovery and replication association analyses for gout vs. asymptomatic hyperuricaemia were conducted in PLINK using an additive logistic regression, with age at recruitment, sex and the first 10 principal components as covariates. Genome-wide associations were annotated using FUMA-SNP2GENE resource. PRS for gout vs. controls were calculated with PRSice v2.0, using the GUGC GWAS summary statistics as the base dataset and the UK Biobank genotype data as the target dataset. PRS for gout vs. asymptomatic hyperuricaemia were generated using the discovery GWAS summary statistics as the base dataset and the replication cohort as the target dataset. Predictive ability of the PRS, the demographic and combined models were assessed using the area under the receiving operating characteristic curve (AUROC).

Results: *Gene expression profiling: BIRC2, CD40LG, CXCL1, IL-1β, MEFV, NLRP12, PANX1, TNFSF14, TXNIP* and *XIAP* showed a significant upregulation in participants with hyperuricaemia with asymptomatic MSU crystal deposits, compared to participants with normouricaemia. *CFLAR, NAIP, NFBIA, NLRC4, NLRP6* and *TLR2* were downregulated in participants during the intercritical stage, compared to the gout flare; however, these differences were not significant after correcting for multiple testing. When these 16 genes were compared among the whole spectrum of normouricaemia to acute gout, *CD40LG, PANX1* and *TNFSF14* showed a downregulation in participants with acute gout, compared to the groups of participants with asymptomatic hyperuricaemia.

Cytokine profiling: In the comparison of intercritical gout vs. gout flare, the levels of VEGF-α and hsCRP showed significant differences. Cytokine levels did not show significant differences among participants with normouricaemia, hyperuricaemia without MSU crystals and hyperuricaemia with asymptomatic deposits of MSU crystals. When cytokines were compared among all groups, GRO-α, IL-1β, IL-6, IL-8, IP-10, MCP-1, TNF-α and hsCRP were greater in the intercritical gout group, compared to the normouricaemia and asymptomatic hyperuricaemia groups.

PRS for gout: The best-fit PRS was generated from 10 SNPs, and showed a significant association with gout. The mean PRS for gout cases and controls were significantly different (0.016 compared to 0.0019, respectively). The predictive ability of the PRS alone was of 62% that increased to a 75% when added to the demographics model.

GWAS and PRS for gout vs. asymptomatic hyperuricaemia: The GWAS revealed 13 independent SNPs located in or near ABCG2, SLC2A9, SLC22A11,

GCKR, MEPE, ADH1B, and the non-coding regions PPM1K-DT and LOC105377323 These loci associated with the transition from asymptomatic hyperuricaemia to gout, and replicated successfully. The PRS generated to distinguish gout from asymptomatic hyperuricaemia was generated from association data of 17 SNPs and gave a predictive ability of 58.2% and 69.2% when combined with the demographics model.

Conclusions: Inflammasome-associated gene expression and cytokine measurements suggest the activation of immune mechanisms in people with asymptomatic deposition of MSU crystals and subclinical systemic inflammation during intercritical gout. The GWAS findings revealed novel loci associated with gout, and confirms the importance of urate transporters and metabolic genes in SU variation and its central role in the pathogenesis of gout. Validation studies in independent datasets are required to confirm gene expression and genome-wide genotype results.

Acknowledgements

…Los fuertes vientos seguirán corriendo, pero tú los recibirás de frente y de pie.

Carlos Plata Vargas (1933-2014)

I would like to start by thanking my supervisors Prof Abhishek, Prof Kevin Morgan and Dr Tamar. Thanks for your patience, your training and the always needed criticism and feedback. Thank you for letting me be independent, but at the same time for always being there to guide me and teach me. I could not have done this without your invaluable support. I am especially grateful to Tamar, you have not only been my guide in the lab, but you have also been a friend every time I needed to talk to someone…¡Gracias!

All my gratitude to the Mexican Council of Science and Technology for funding this PhD project and to the RoseTrees charity for providing financial support for consumables. I would like to express my gratitude to all the participants who made this research possible, those who took part of the clinical studies at the Department of Academic Rheumatology and the UK Biobank project.

Thanks to all the members in the Human Genetics Lab and the Department of Academic Rheumatology. Special thanks to Dr Keeley Brookes for always being willing to help and advise. Huge thanks to Dr Sally Chappell, having you as a mentor during my MSc project inspired me not only to start this PhD journey, but also to know the type of professional I want to become.

Massive thanks to my housemates and friends: Jahz, David, Beto, Ana, Blanca, Lu, Gaby, Jess, Edu and Tesh. Thanks to you, I learned that family isn't always blood. Beyond the professional achievements, this PhD experience has already been worth it for having met such wonderful people as you all. I would like to

thank Jahz, Jess and Gaby, having you around during the craziness of the last year has been an enormous blessing. Jahz: thank you for being since day one to listen to all the personal and academic rant, for the shared cravings, glasses of wine, ciders, walks, trips, tv shows, movies, and countless cups of coffee and tea. If anyone can understand the road of these years it's you, and I could not have had a better friend to share them with.

Thanks also to my friends in Mexico, for the quick visits and endless texts. Because even thousands of kilometres away and having your own activities and busy lives, you always had time for a call or a text. And believe me, in several occasions they were the motivation I needed to keep working and surviving this roller coaster. I am especially grateful to my friend and mentor Gabriel Sandoval, your time, advice, and guidance have been an inspiration to keep doing science. Special thanks to Silvia and Badu, you have helped me in imaginable ways throughout my life. Badu, this thesis is dedicated to your dad too, with all my love and admiration.

Thanks to all my family –uncles, cousins, nephews. Because physical distance only strengthened bonds that were already powerful. Thank you for not only being aware of my adventures and professional growth, but also for being a key support to my parents during these years. Special thanks go to my female cousins, for being 24/7 no matter if it involved suffering or celebrating together. In Mexico, women represent only 33% in science community; gender gap is a sad reality in most areas. Mony, Cecy, Mariana, Raquel, Pau Plata, Itxel, Gis, Frida, Pau Sandoval, Adriana, Cons, Lupi, MaJo, Romina, Victoria, Mía and Kass do not let anyone tell you "you can't do something". Dream big and work hard, everything is possible, always.

Thank you Claudio, you came into my life to brighten up my days. I could not have gone through the last months of this journey without your patience, support, company, jokes and love. Grazie mille amore!

Massive thanks to my grandparents, Carlos, Pablo, Rafaela and Consuelo, your example of hard work, strength and endless love has been crucial to become the woman I am today.

Thanks to my brother Alfredo, for always being such an inspiration and example of passion, humility and kindness. Thanks for your support, for believing in me and for always being my fan number one. I could not have asked for a better partner in crime.

Finally, massive thanks to my parents Gaby and Alfredo. For never minimising my dreams, for letting me be curious and creative, for raising an empowered and independent girl. Words would never be enough to describe the love and gratitude I feel, nor would they repay your efforts and sacrifices that allow me to be where I am at today. Thank you for giving me these immense wings and being the wind beneath them. ¡Los amo!

This thesis is dedicated to my grandparents Carlos, Pablo, Rafaela and Consuelo.
Guerreros incansables y amantes de la vida.
Don Carlitos y Rafita, los extraño cada segundo. Abrazos hasta el cielo.

Agradecimientos

…Los fuertes vientos seguirán corriendo, pero tú los recibirás de frente y de pie.

Carlos Plata Vargas (1933-2014)

Me gustaría empezar agradeciendo a mis supervisores, Prof Abhishek, Prof Kevin y Dr Tamar. Por su paciencia, disposición y crítica siempre necesaria para mi crecimiento profesional. Gracias por dejarme ser independiente, pero por siempre estar para guiarme y enseñarme. Gracias en especial a Tamar, no solo por haber sido mi guía en el laboratorio, sino por haber sido una amiga y escucharme siempre que había momentos complicados.

Todo mi agradecimiento al Consejo Nacional de Ciencia y Tecnología (CONACYT) en México por otorgarme financiamiento para realizar este Doctorado. Gracias también a RoseTrees por proveer con fondos adicionales. Mi total agradecimiento hacia todos los participantes que hicieron posible la realización de este proyecto; a todos los que formaron parte de los estudios clínicos del Departamento de Academic Rheumatology y del UK Biobank.

Gracias a todos los miembros de Human Genetics y Academic Rheumatology. Gracias especialmente a Dr. Keeley por tu disposición siempre que pedía tu ayuda. Mi profundo agradecimiento a Dr. Sally Chappell, tu guía y enseñanza durante la Maestría no solo me motivaron a iniciar el Doctorado, sino que además me mostraron el tipo de profesionista que quiero ser.

Gracias a mis ex-compañeros de casa y amigos: Jahz, David, Beto, Blanca, Lu, Ana, Gaby, Jess, Edu y Tesh. Cada uno de ustedes es la prueba ferviente de que la familia no necesariamente es de sangre. Un agradecimiento especial a

Jahz, Gaby y Jess: haberlas tenido de compañeras de pandemia para sobrevivir la locura de este último año ha sido una enorme bendición. Jahz: gracias por estar desde el día uno para escuchar la locura tanto académica como personal, por los antojos compartidos, copas de vino, caminatas, viajes, e incontables tazas de café y té. Porque si alguien puede entender el caminar de estos años eres tú y no pude haberlo compartido con mejor compañera. Más allá de los logros profesionales, la experiencia del doctorado ya ha valido la pena por haber conocido a personas tan maravillosas como todos ustedes. ¡Gracias!

Gracias a mis amigos en México, por las visitas exprés y mensajes. Porque aún a miles de kilómetros y con sus ocupaciones múltiples, siempre tuvieron tiempo para una llamada o mensaje, que créanme, en diversas ocasiones fue mi motivación para seguir resistiendo esta montaña rusa. Agradecimientos especiales a Gabriel Sandoval, tu tiempo para escuchar, aconsejar y enseñar ha sido inspiración para seguir haciendo ciencia. Gracias especialmente a Silvia y Badu, su apoyo me ha inspirado y motivado de maneras que no se imaginan a lo largo de mi vida. Badu, esta tesis va con dedicatoria especial a tu papá también, con toda mi admiración y cariño.

Gracias a toda mi familia, tíos, primos y sobrinos, porque la distancia física solo fortaleció lazos que ya eran fuertes. Gracias por no solo estar al pendiente de mis aventuras y crecimiento profesional, sino además por ser soporte para mis papás durante estos años. Gracias especiales a mis primas, por estar 24/7 para sufrir o gozar conmigo. En México, las mujeres en la ciencia ocupan tan solo el 33%. La brecha de género es una realidad en todos los ámbitos. Mony, Ana Cecy, Mariana, Raque, Pau Plata, Itxel, Gis, Frida, Pau Sandoval, Lupis, Adri,

Cons, MaJo, Romina, Victoria, Mía y Kass, que nadie les diga nunca que no pueden, sueñen grande y trabajen duro. Todo se puede, siempre.

Gracias Claudio por llegar a alegrar mis días. No pude haber pasado estos últimos meses sin tu apoyo, tu paciencia, tu compañía, tus bromas y tu amor. Grazie mille amore!

Gracias a mis abuelos, Carlos, Pablo, Rafaela y Consuelo, su ejemplo de lucha, trabajo y amor incondicional ha sido determinante para ser la mujer y profesionista que hoy soy.

Gracias a mi hermano Alfre, por ser una gran inspiración de pasión, humildad y amabilidad. Gracias por todo tu apoyo, por creer en mí y por ser mi fan número uno siempre. No pude haber pedido un mejor compañero de aventuras.

Gracias a mis padres, Gaby y Alfredo, por nunca minimizar mis sueños, por dejarme ser curiosa y creativa, por criar a una niña empoderada e independiente. Las palabras que pueda plasmar en esta tesis, nunca serían suficientes para describir el amor y agradecimiento que les tengo, ni para retribuir sus esfuerzos y sacrificios que hoy me permiten estar en este lugar. Gracias por darme unas alas inmensas y ser el viento que las impulsa. Gracias, gracias, gracias. ¡Los amo!

Esta tesis está dedicada a mis abuelos Carlos, Pablo, Rafaela y Consuelo.
Guerreros incansables y amantes de la vida.
Don Carlitos y Rafita, los extraño cada segundo. Abrazos hasta el cielo.

List of publications

Published papers

- Sandoval-Plata G, Morgan K, Abhishek A. Variants in urate transporters, ADH1B, GCKR and MEPE genes associate with transition from asymptomatic hyperuricaemia to gout: results of the first gout versus asymptomatic hyperuricaemia GWAS in Caucasians using data from the UK Biobank. Ann Rheum Dis. 2021 Sep;80 (9):1220-1226. doi: 10.1136/annrheumdis-2020-219796. Epub 2021 Apr 8. PMID: 33832965.

Presentations at Conferences

- American College of Rheumatology Annual Meeting (Atlanta, 2019): Asymptomatic Monosodium Urate Crystal Deposition Associates with Increased Expression of Pro-Inflammatory Genes.

Publications related to the thesis

- Sandoval-Plata G, Nakafero G, Chakravorty M, Morgan K, Abhishek A. Association between serum urate, gout and comorbidities: a case-control study using data from the UK Biobank. Rheumatology (Oxford). 2021 Jul 1;60(7):3243-3251. doi: 10.1093/rheumatology/keaa773. PMID: 33313843.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABCG2: ATP Binding Cassette Subfamily G member 2
aggNETs: Aggregated neutrophil extracellular traps
ALDH: Aldehyde dehydrogenase
AMP: Adenosine monophosphate
ANOVA: Analysis of Variance
ASC: Apoptosis speck protein
AUROC: Area under the receiver operating characteristics
BMI: Body mass index
bp: Base pair position
BIR: Baculovirus inhibitor of apoptosis domain
CARD: Caspase activation and recruitment domain
cDNA: Complementary DNA
CKD: Chronic kidney disease
CPPD: Calcium pyrophosphate dihydrate
C_T: Threshold cycle
Chr: Chromosome
CHARGE: Cohort for Heart and Aging Research in Genome Epidemiology
DAMPs: Damage-associated molecular patterns
DECT: Dual-energy computer tomography
DNA: Deoxyribonucleic acid
EULAR: European League Against Rheumatism
FUEA: Fractional excretion of uric acid
FDR: False discovery rate
FUMA-GWAS: Functional Mapping and Annotation of Genome Wide Association Studies
gDNA: Genomic DNA
GMP: Guanosine monophosphate
GOI: Gene of interest
GUGC: Global urate genetics consortium
GWAS: Genome wide association analysis
hsCRP: High sensitive C-reactive protein
HPRT: Hypoxanthine-guanine phosphoribosyltransferase
HWE: Hardy-Weinberg equilibrium

IMP: Inosine monophosphate

LD: Linkage disequilibrium

LRR: Leucine-reach repeat

mRNA: Messenger RNA

MSD: Meso scale discovery

MSU: Monosodium urate

NETs: Neutrophils extracellular traps

NICE: National Institute for Health and Care Excellence

NHANES: National Health and Nutrition Examination Survey

NLR: Nucleotide-like receptors

NLRP3: NLR family pyrin domain containing 3

NOD: Nucleotide oligomerisation domain

NSAIDs: Nonsteroidal anti-inflammatory drugs

NTC: Non template control

OR: Odds ratios

PAMPs: Pathogen-associated molecular patterns

PAR: Pseudo autosomal regions

PCs: Principal components

PCA: Principal components analysis

PRPP: Phosphoribosyl pyrophosphate

PRRs: Pattern recognition receptors

PRS: Polygenic risk scores

PYD: Pyrin domain

QC: Quality control

RIN: RNA integrity number

RNA: Ribonucleic acid

ROL: Renal overload

ROS: Reactive oxygen species

RUE: Renal underexcretion

SD: Standard deviation

SE: Standard errors

SLC-A-: Solute carrier family – member -

SNP: Single nucleotide polymorphism

SOG: Sons of gout

SOG-AS: Switch off gout attack study

SU: Serum urate

SUA: Serum uric acid

TLRs: Toll-like receptors

TRX: Thioredoxin

TXNIP: TRX-interacting protein

UA: Uric acid

ULT: Urate lowering treatment

URAT1: Urate transporter 1

UUE: Urinary urate excretion

XOI: Xanthine oxidase inhibitors

Chapter 1. Introduction

Gout is the most common form of inflammatory arthritis, characterised by episodes of acute inflammation because of shedding of monosodium urate (MSU) crystals deposited within and around joints. Crystal deposition occurs when the serum urate (SU) increases beyond its solubilisation point of 6.75 mg/dL at 37ºC, according to *in vitro* studies (Nicola Dalbeth, Merriman, & Stamp, 2016; Loeb, 1972). Although studies have shown that at temperatures between 30-35ºC (the average temperature of the first metatarsophalangeal joint in humans), the saturation point decreases to between 4.5 and 6.0 mg/dL (Nardin, Fogerson, Nie, & Rutkove, 2010).

Gout has a prevalence of 2.49% in the United Kingdom, and is more common in men (with a prevalence between 3 to 6%) than in women (prevalence of 2%) (Kuo, Grainge, Mallen, Zhang, & Doherty, 2015). Besides gender, age is another factor that contributes to the development of gout, increasing the prevalence up to 9% in men and 6% in women older than 75 years of age (H. K. Choi, Mount, & Reginato, 2005; K. L. Wallace, Riedel, Joseph-Ridge, & Wortmann, 2004; Zhu, Pandya, & Choi, 2011). If gout is left untreated, it can cause tophi and evolve to chronic arthritis in some but not all patients. This generates greater costs to the health services, due to an increased use of the healthcare resources (Saseen et al., 2012). In addition, patients with gout often develop other comorbidities that not only aggravate health related quality of life, but also hinders an adequate management and treatment of the disease. For instance, chronic kidney disease limits the maximum allopurinol dosing that can be used (Drivelegka, Sigurdardottir, Svärd, Jacobsson, & Dehlin, 2018).

Studies have focused on understanding the inflammatory response to MSU crystals, the mechanisms involved in the resolution of flares, and the risk factors associated to SU concentrations and gout. This chapter provides relevant information about the nature and historical background of gout, the clinical characteristics and diagnosis, the metabolic mechanisms of hyperuricaemia, the immune response during gout flares, the genetic basis in the pathogenesis of the disease, and finally the rationale behind this study concept.

1.1 History of gout

Gout is one of the rheumatic diseases that has been described throughout centuries. Its history dates to ancient times, and thanks to anthropologic discoveries, it is known that Tyrannosaurus rex fossils showed evidence of bone erosion characteristic of gout (Rothschild, Tanke, & Carpenter, 1997). Additionally, skeleton remains from Egypt and England (~2480 BC and 150 AD respectively), presented similar bone damage features because of gouty erosions. However, first documented reports of this disease track back to the Ancient Greece and the Roman Empire. Hippocrates first described several gout epidemiological facts, such as the higher prevalence in men and postmenopausal women, and he assigned the word “podagra” to describe severe swelling and pain in the foot (SydenhamSociety, 1891). The origin of the word “gout” (from the Latin word “gutta”, which means drop) is credited to Ralph Bockingus, who based on the Hippocratic theory of the four biologic humours, attributed the infiltration of a biologic fluid into the joint as the causal agent of articular inflammation (S. L. Wallace, 1964).

For many centuries, gout was considered to be an upper social class-exclusive disease, due to its popularity among aristocrats, whose lifestyles were characterised by excessive food and alcoholic drinks intake. However, it was not until the second century AD, when gout’s hereditary nature was introduced after observing its incidence among generations. During this period, gout was also defined as a disease with intermittent symptoms (S. L. Wallace, 1964). The sixth century AD marked an additional and crucial discovery for gout, when *Colchicum autumnale* extracts were introduced as the first therapy to treat gout (Storck, 1764).

Modern history in gout understanding is marked by the English physician Thomas Sydenham. He provided a detailed and accurate description of symptoms during gout flares, as well as several clinical features typical of the chronic stage, such as joint destruction, tophi formation and renal stones (Sydenham, 1683). In 1776, the chemists Carl Wilhelm isolated an acid substance from renal stones. Twelve years after, William Hyde examined a tophus, concluding the acid substance previously described by Wilhelm, was also present in the tophus. This acid compound was later identified as uric acid, and in 1848 Alfred Baring Garrod observed that patients with gout had high levels of uric acid and gout flares could be the consequence of uric acid precipitation into joint tissues (Garrod, 1848).

One of the breakthrough discoveries happened in the 1960s when McCarthy and Hollander observed MSU crystals in synovial fluid of patients with gout using polarised light microscopy. This discovery allowed them to conclude that episodes of inflammation were induced by the presence of MSU crystals (McCarty & Hollander, 1961). This technique has been established as the gold standard for the diagnosis of gout and calcium pyrophosphate dihydrate (CPPD) crystal deposition disease.

1.2 Nature of gout

After the historical observations, it was recognised that high concentration of SU and its crystallisation within joints were the causal agents of gout. However, it was not until the 20th century when researchers observed that after injecting MSU crystals, the surrounding cells phagocytised them and initiated a severe inflammatory response. Thanks to these discoveries, it is now known that the natural history of gout is defined by four stages (Nicola Dalbeth et al., 2016):

- Hyperuricemia without MSU crystal deposition
- Hyperuricemia with MSU crystal deposition but no flares or tophi
- MSU crystal deposition triggering episodes of acute inflammation (gout flares)
- Chronic tophaceous gout

Hyperuricaemia is defined as SU levels greater than 6.8 mg/dL [408 µmol/L], concentration at which urate precipitates *in vitro* under physiological conditions (37ºC and pH 7.4) (Loeb, 1972). The mechanisms leading to an overproduction or underexcretion of urate that cause hyperuricaemia, will be described in detail in section 1.6.1.

Urate crystal formation is the key stage between hyperuricaemia and gout; however, the relationship between SU levels and MSU crystallisation is complex. Although SU concentration is important, additional factors contribute to the process of microcrystals formation, also known as nucleation. Among them, there are several physicochemical factors such as temperature, pH, and movement or mechanical trauma in the joints (Loeb, 1972; Wilcox & Khalaf, 1975). A study reported that different components in the cartilage and the synovial fluid are also essential for nucleation and growth of MSU crystals. This was determined after analysing nucleation in synovial fluid of patients with

different rheumatic diseases, and observing that gouty joints promoted an enhancement in crystal formation that occurred through mechanisms independent of SU levels (Martillo, Nazzal, & Crittenden, 2014; McGill & Dieppe, 1991).

The immune response to MSU crystals is the fundamental mechanism behind a gout flare. Briefly, this response is mediated by the NLR family Pyrin Domain Containing 3 (NLRP3) and the Toll-like Receptors (TLRs) 2 and 4, that together promote the synthesis of pro-inflammatory cytokines that amplify the immune response (Shi, Mucsi Ashley, & Ng, 2009). A detailed explanation of this will be provided in section 1.7. Gout flares self-resolve a few days to up to 2 weeks after the episode of acute inflammation started (H. K. Choi et al., 2005). The resolution phase is mediated by the formation of aggregated neutrophil extracellular traps (aggNETs) that inactivate the synthesis of pro-inflammatory molecules (Schett, Schauer, Hoffmann, & Herrmann, 2015). If hyperuricaemia persists and gout is not managed accordingly, gout flares may occur more frequently and affect multiple joints.

The time of evolution from acute gout to chronic tophaceous gout is highly variable among patients. However, the continuous deposition of MSU crystals within joints leads to the development of tophus. These are deposits of densely compacted MSU crystals and inflammatory cells (macrophages and plasma B cells) surrounded by a fibrovascular capsule. Tophus can be present without symptoms of local inflammation or present with flares. However, as they continually stimulate the secretion of cytokines and proteases by interacting with osteoblasts and chondrocytes, tophi can cause bone erosion, fibromatosis and cartilage loss (Chhana & Dalbeth, 2015). Advanced gout manifests with frequent flares and a greater number of tophi. This stage of the disease is

associated with a decline of the physical and mental health of the patient, affecting daily activities and work productivity (Khanna et al., 2012). In addition, patients with gout often develop other comorbidities that not only aggravate health related quality of life, but also hinder an adequate management and treatment of the disease (Drivelegka et al., 2018).

1.3 Signs and Symptoms of gout

The hallmark sign of gout is an abrupt episode of inflammation in a joint. Upon examination during gout flares, patients present severe pain, redness, stiffness, sensitivity to touch and warmth in the affected joint. Mild fever and fatigue may be present as indicators of systemic inflammation. During the initial stages, gout flares are usually monoarticular and affect more frequently the first metatarsophalangeal joint, followed by the ankle, transverse tarsal joint, knee, wrist and elbow (H. K. Choi et al., 2005). Flares usually occur at night, presumably due to a decrease of the body temperature that facilitate MSU crystallisation, nocturnal dehydration, variation of cortisol levels and sleep apnoea that enhance purine metabolism via nucleotides turnover (Hyon K. Choi et al., 2015). Gout flares can be triggered by dehydration, surgical procedures, alcohol consumption (especially beer) and the intake of purine rich food (beef, lamb and seafood) (A. Abhishek, Valdes, Jenkins, Zhang, & Doherty, 2017; Rothenbacher, Primatesta, Ferreira, Cea-Soriano, & Rodríguez, 2011). When hyperuricemia persists or patients do not receive any treatment, gout becomes more severe and flares can occur more frequently and manifest as polyarticular inflammation or as chronic tophaceous gout (Nicola Dalbeth et al., 2016; Parathithasan, Lee, Pianta, Oon, & Perera, 2016). In patients with advanced gout, the principal sign is the presence of tophi in metatarsophalangeal joints, metacarpophalangeal joints, elbows, ankles or ears. Tophi tend to appear many years after the first flare if hyperuricaemia continues and after recurrent episodes of inflammation. However, some patients develop aggressive tophaceous gout within a few years, typically in people with organ transplants treated with cyclosporine (Clive, 2000). And although they are often clinically silent, long-term, they can cause extreme tenderness, limited movement of the affected joints, skin ulceration and infection, and psychological impact owing to

restrictions in social and work activities and the cosmetic effects (Chhana & Dalbeth, 2015).

1.4 Diagnosis of gout

There are multiple guidelines published by different Rheumatology societies and organisations that establish the criteria to diagnose and classify gout. Among these are the American College of Rheumatology (ACR) and the European League Against Rheumatism (EULAR) (Vargas-Santos, Taylor, & Neogi, 2016). In 2015, the ACR and EULAR published a collaborative document consisting of improved criteria that consider several observations, such as clinical features of gout (joint involvement, characteristics of the affected joint, duration of the episode and presence of tophus), laboratory results (SU levels and presence of MSU crystals) and imaging (radiography, ultrasound and double-energy computed tomography). The sensitivity and specificity were 92% and 89% respectively, representing the best values compared to previous published criteria. It is however, recommended exclusively for study purposes and not for clinical practice, where aspiration of synovial fluid remains the gold standard (Neogi et al., 2015). In addition to the ACR/EULAR collaborative guideline, in 2018 EULAR published an updated list of eight recommendations for the diagnosis of gout. Besides arthrocentesis, evaluation of gout-related clinical features and the use of imaging methods, these recommendations stated the importance of conducting a thorough assessment of the risk factors that contribute to chronic hyperuricaemia, such as chronic kidney disease, obesity, specific medication for hypertension and dietary elements (consumption of alcohol, sugar drinks, meat and shellfish), and the screening of comorbidities that coexist with gout, like diabetes, hypertension, hypercholesterolemia, ischaemic heart disease, etc. (Pascal Richette et al., 2019). Therefore, laboratory tests are essential not only to measure SU levels, but also to analyse glycaemic and lipidomic profiles, kidney function test, blood cells count and C-reactive protein measurement, to complete the diagnosis, to

determine potential therapeutic implications and to monitor the progression of the disease or response to therapy (Qing et al., 2013; Ragab, Elshahaly, & Bardin, 2017).

Currently, arthrocentesis of the affected joint to confirm the presence of MSU crystals is the technique that allows a definite diagnosis of gout (Neogi et al., 2015; Pascal Richette et al., 2019). However, joint aspiration is frequently not performed, especially in primary care, where most gout diagnosis are identified, or is inaccessible when flares occur in small joints (Harrold et al., 2013). Furthermore, the gold standard though having a good sensitivity and specificity (69% and 97% respectively) is still susceptible to human error, particularly because it requires qualified personnel and adequate equipment (Gordon, Swan, & Dieppe, 1989). Moreover, gout and other arthritis may coexist, and an adequate diagnosis is vital to select further management (Genes & Chisolm-Straker, 2012). For those reasons, diagnosis of gout should be completed by an exhaustive physical evaluation, a record of medical history, laboratory tests and imaging methods. The following subsections include a brief description of the gold standard and the main imaging methods used for the diagnosis of gout.

1.4.1 Synovial fluid analysis

The aspiration of synovial fluid or tophus to identify MSU crystals under polarising light microscopy is considered as the gold standard diagnostic test. This technique involves the analysis under polarising light microscopy of a droplet of fresh synovial fluid to look for intracellular needle shaped MSU crystals with negative birefringence. MSU crystals show a bright yellow or blue colour, depending on the position of the light compensator, with lengths that

range from 1-20 µm (Figure 1.1). (H. Paul, Reginato, & Schumacher, 1983; Ragab et al., 2017)

Observing the physical characteristics of crystals is essential to distinguish gout from other types of arthritis that not only mimic clinical features of gout, but are also triggered by crystals, such as calcium pyrophosphate dihydrate (CPPD) (responsible for acute CPP crystal arthritis - previously called pseudogout) and basic calcium phosphate (Gordon et al., 1989; MacMullan & McCarthy, 2012). Septic arthritis also presents similar symptoms as gout and should be excluded when analysing the synovial fluid by gram staining and culturing (Ragab et al., 2017). Apart from looking for MSU crystals, the analysis of synovial fluid is important to determine cell counts. During gout flares, the number of leukocytes varies around 50,000 cells/µL; more than 50,000 cells/µL with a higher percentage of neutrophils is indicative of septic arthritis (Pascual, Batlle-Gualda, Martinez, Rosas, & Vela, 1999; Ragab et al., 2017).

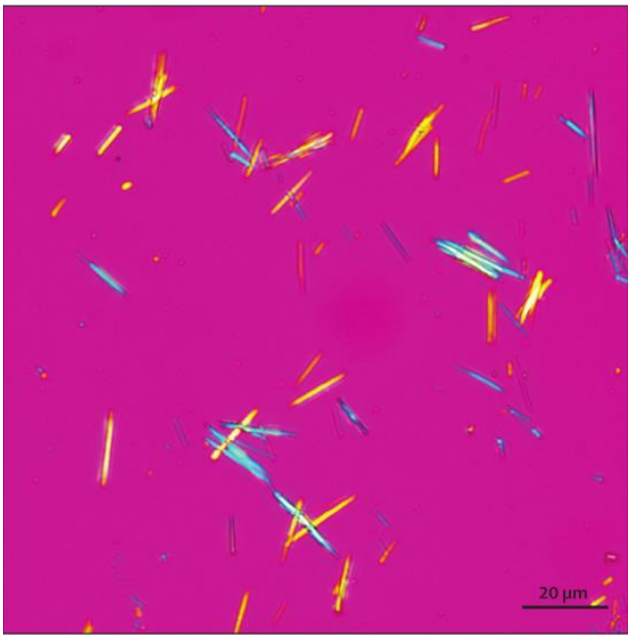


Figure 1.1 Monosodium urate crystals under polarising light microscopy.
MSU crystals show a needle shape with negative birefringence. Figure taken from (Nicola Dalbeth et al., 2016).

1.4.2 Imaging in gout

A wide variety of imaging methods including plain radiographs, ultrasound, conventional tomography, dual-energy computer tomography (DECT), and magnetic resonance, have been introduced as potential tools to identify patients with gout, to monitor the progression of the disease and to assess the response to treatment (Nicola Dalbeth & Doyle, 2012; Parathithasan et al., 2016). Among these imaging tools, plain radiographs, DECT and ultrasound are the most used. Each one presents advantages and disadvantages, and using one instead of another depends on several factors, for example the stage of gout during examination, availability in clinical settings, expertise of the health professionals, etc. Imaging has acquired more interest as some of the tools are useful in guiding and facilitating synovial fluid aspiration, and they represent a potential alternative to detect joint abnormalities when they are not clinically detectable as no tophi are present or there are no evident signs of acute inflammation (during asymptomatic hyperuricemia or intercritical gout) (N. Dalbeth et al., 2015; Naredo et al., 2014).

1.4.2.1 Plain radiography

Given that radiographs are widely available and usually inexpensive, they have been used for a long time in the diagnosis of gout. The radiographic evaluation in gout patients allows the observation of different non-specific features that occur because of an acute flare (such as soft-tissue swelling), but more specifically changes of chronic tophaceous gout that occur after sustained deposition of MSU crystals and its evolution to chronic tophaceous gout. In early gout, radiographs are typically normal, but swelling of soft tissues might be visible (Nicola Dalbeth & Doyle, 2012). In advanced gout, tophi can be observed

as dense masses between soft tissues and bones and if calcified, can be indicative of renal failure causing abnormalities in calcium metabolism or simply long-term urate deposits on which calcification has occurred (Gentili, 2003). Articular and periarticular bone erosions are also typical in patients with chronic tophaceous gout. On plain radiographs, these erosions present a characteristic sclerotic border and overhanging edges (Nicola Dalbeth & Doyle, 2012). Additionally, intraosseous tophi can cause calcifications that are also visible on radiographs. Figure 1.2 shows some examples of these type of radiographic features. Provided that bone and cartilage damage appear in advanced gout, usually after 5-10 years or even longer after the onset of the disease, plain radiographs are not useful in the diagnosis during early stages but assist in monitoring the progression of gout. In addition, they help in providing a differential diagnosis. For instance, in early gout the joint space keeps a normal width, while other arthropathies such as rheumatoid arthritis and osteoarthritis show a narrow articular space, although this might change in advanced gout (Gentili, 2003).

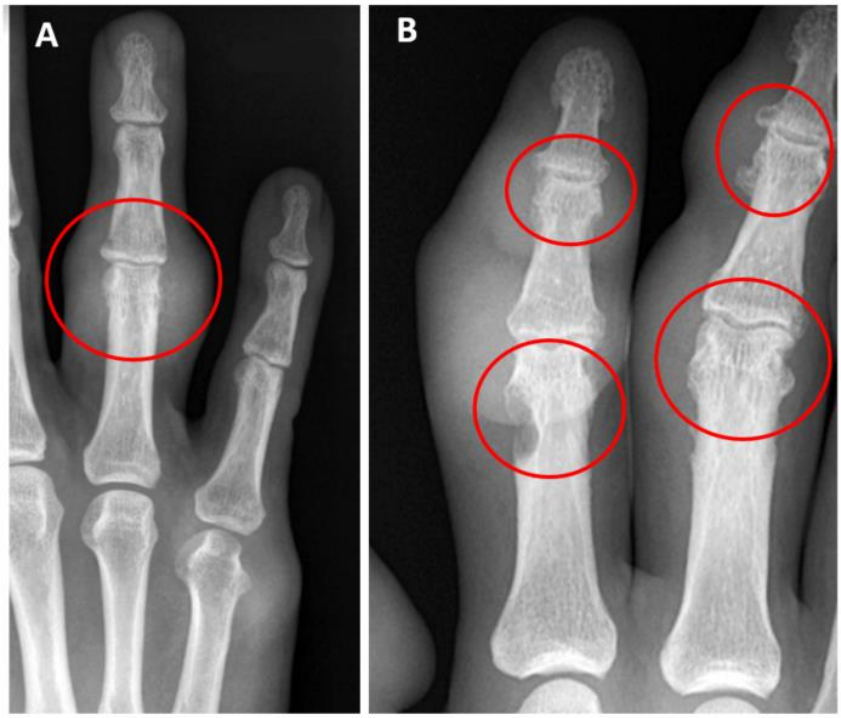


Figure 1.2. Dorsopalmar radiographs.
Showing a dense mass around soft tissues (A) and bone erosion with the characteristic overhanging edges caused by large tophi in the index and middle fingers (B). Adapted from (Gentili, 2003)

1.4.2.2 Ultrasonography

Even though compared to plain radiographs, ultrasound may not be necessarily accessible and requires specialised personnel, it provides a more exhaustive evaluation and is not invasive (Filippucci, Reginato, & Thiele, 2020). Ultrasonography allows the observation of more parameters distinctive of gout such as MSU crystal deposits, bone erosion, etc. during early and advanced gout. The most distinguishing sign observed by ultrasound is known as “the double contour sign” that forms because of the hyperechoic projection of urate deposits around the hyaline cartilage that can be visualised as a line of a thickness like or even greater than that of the bone (Figure 1.3). The double contour sign is not attributable to other crystal-induced diseases and has a high specificity (>90%), which makes it a good alternative when arthrocentesis is not

accessible. However, it has shown a variable sensitivity (20-90%) (Thiele & Schlesinger, 2007). After the increasing use of ultrasonography in rheumatology practice and the emergence of more studies, there is now a better consensus of ultrasound definitions for gout (Gutierrez et al., 2015). In that way, synovitis during a gout flare (with joint effusion and Doppler signal indicating hypervascularity), bone erosion, tophi and hyperechoic aggregates, are some other ultrasonographic findings of gout (Nicola Dalbeth & Doyle, 2012).

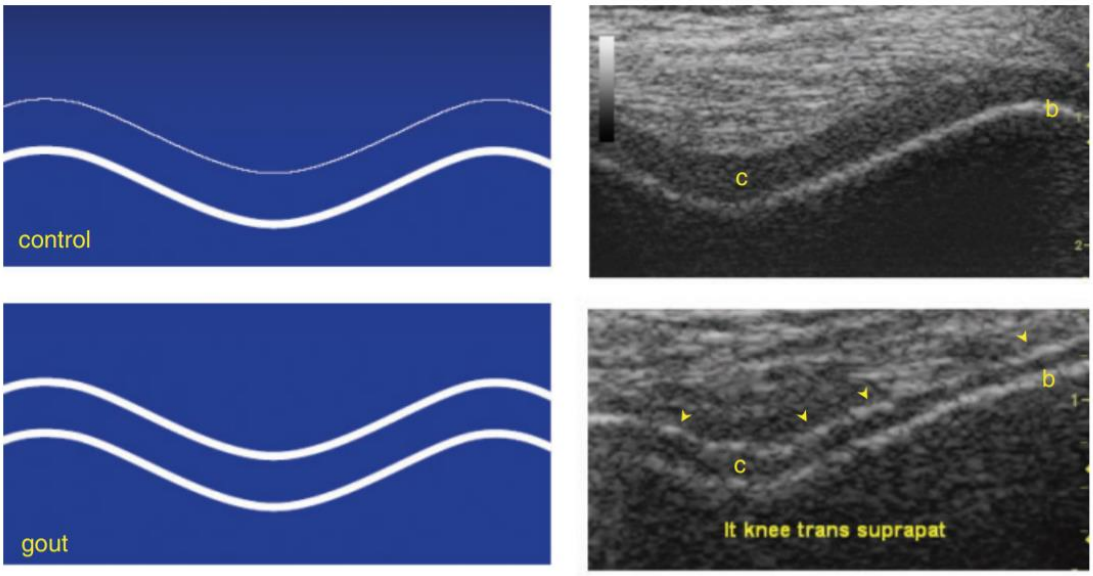


Figure 1.3. Comparison of ultrasound images of healthy vs gouty knee joints showing the double contour sign.
Taken from (Thiele & Schlesinger, 2007).

1.4.2.3 Dual Energy Computed Tomography (DECT)

Similarly to ultrasound, DECT is useful in identifying features during the initial stages of gout. It is particularly valuable to distinguish among different arthritis and to diagnose subclinical gout cases (i.e. presence of tophi without evident clinical manifestations) (Gentili, 2003). This is one of the most accurate methods in detecting MSU crystals and differentiating them from calcium deposits or

other type of crystals thanks to their different x-ray absorption characteristics. DECT mechanism consists of placing two x-ray detectors (usually at 80kVp and 140 kVp) that generate simultaneous images at different energy levels, these are then converted into 3D graphics with colour coding according to the dual energy properties (McQueen, Doyle, & Dalbeth, 2011). Figure 1.4 shows an example of a DECT scan of the foot with multiple tophi. Compared to other methods, DECT provides a clearer picture of the tophi sizes and bone and cartilage involvement. However, like ultrasound, it has a high specificity (up to 95%) but a lower sensitivity that is highly dependent on the stage of gout under study. Additionally, as it uses radiation, it still carries risk associated with radiation (e.g. increased risk of cancer) and is not that accessible in terms of costs and availability in clinical practices (Bayat, Baraf, & Rech, 2018; McQueen et al., 2011).

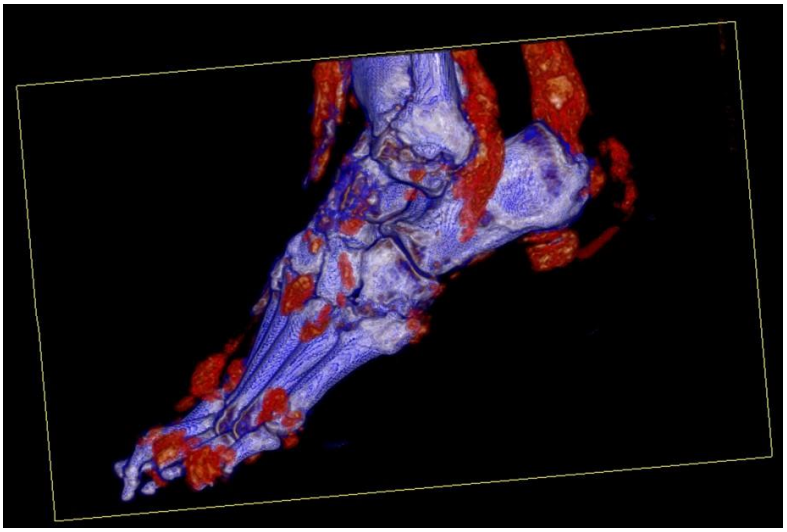


Figure 1.4 Dual energy computed tomography scan.
Red spots represent tophaceous material around bone and soft tissue of different sites of the foot. Taken from (McQueen et al., 2011)

1.5 Management of gout

As for the classification, the ACR (FitzGerald et al., 2020), the EULAR (P Richette et al., 2017) and the British Society for Rheumatology (Hui et al., 2017) have published guidelines for the management and treatment of gout. These recommendations divide the management in non-pharmacological and pharmacological treatment.

Non-pharmacological treatment includes changes in lifestyle to reduce the number of flares and prevent the development of other metabolic disorders. Dietary modifications such as reduced intake of meat and seafood, limited consumption of drinks with high content of fructose (carbonated beverages and energy drinks) or alcoholic drinks (especially beer, wine and spirits), and increased ingestion of low caloric food, are recommended to reduce the risk of gout occurrence (Yuqing Zhang et al., 2012; Yuqing Zhang et al., 2006). A study showed that a weight loss of 7.7 kg, after 16 weeks of calorie restriction diet, had a meaningful effect in lowering SU concentrations of 0.47-0.57 mmol/L, and decreasing frequency of flares (Dessein, Shipton, Stanwix, Joffe, & Ramokgadi, 2000). Moreover, regular exercise has also shown an impact in reducing SU levels. According to a study, running more than 8 km per day is associated to a gradual decline in SU and lower risk (50%) of developing gout (Williams, 2008). Non pharmacological interventions have small effect on serum urate.

On the other hand, pharmacological treatment depends on the stage of the disease, but in general, the management focuses on three targets (Schlesinger, 2004):

- Treating a gout flare
- Decreasing SU levels to dissolve crystals
- Providing prophylactic treatment to avoid future flares

1.5.1 Gout flares management

Treatment during gout flares should focus on decreasing inflammation and pain. Even though the ACR and the EULAR gout treatment guidelines do not clearly indicate when to start with urate lowering treatment (ULT) in relation to a gout flare (Q. Li et al., 2017), it has been advised that an abrupt variation in SU levels can lead to a partial dissolution of urate crystals, resulting in their mobilisation and consequent flare aggravation (H. K. Choi et al., 2005).

Among the most common drugs used to treat flares, are the nonsteroidal anti-inflammatory drugs (NSAIDs), colchicine, corticosteroids and IL-1β blockers. The selection of treatment depends on clinical features and medical history of each patient. Colchicine and NSAIDs (especially naproxen, ibuprofen and diclofenac) are commonly used and are effective for flares when administered in the appropriate doses (Roddy, Mallen, Hider, & Jordan, 2010). However, as they are metabolised in liver and kidney, they are contraindicated in patients with chronic kidney disease, cirrhosis or patients undergoing treatments that inhibit P-glycoprotein or CYP3A4 (P Richette et al., 2017; Schlesinger, 2004).

Additionally, they are contraindicated in people with ischaemic heart disease, congestive cardiac failure and many other comorbidities as outlined in the British National Formulary. In these cases, oral, intra-articular or intramuscular administration of corticosteroids can be used (Jansen & Rasker, 2011). Anti-IL-1 monoclonal antibodies canakinumab or anakinra, are prescribed when patients show intolerance to colchicine, NSAIDs and corticosteroids (Bardin, 2015). Application of a single dose of canakinumab has proven to be effective in relieving pain 24 hours after its application, and to extend the period without flares (Schlesinger et al., 2012). Despite the promising results, the use of

monoclonal antibodies is not the first choice in clinical practice, due to its high cost in the UK and elsewhere.

1.5.2 Intercritical gout management

There are two categories of ULT: xanthine oxidase inhibitors (XOI) and uricosurics. Among XOI, allopurinol and febuxostat are considered as first and second line treatments, respectively. In cases where there are contraindications (especially in patients with severe renal failure) or low response to this type of treatment, uricosurics are therefore prescribed. Probenecid, sulfinpyrazone, benzbromarone and lesinurad are the most commonly used uricosurics. Severe cases of gout with little or no response to XOI or uricosurics, are treated with pegloticase. This enzyme converts uric acid into a more soluble metabolite. However, the options of treating gout patients with these medicines is limited as pegloticase is not available in the UK, benzobromarone is not licensed for the use of gout in the UK and is only available from Europe on a named patient basis only, and lesinurad does not have NICE (The National Institute for Health and Care Excellence) approval and has been recently withdrawn from different markets by the manufacturer. (FitzGerald et al., 2020; P Richette et al., 2017)

There is a consensus about prescribing ULT at low doses initially, and increase it gradually based on kidney function tests and SU levels. Losartan or calcium channel blockers should be used in patients with hypertension (Hyon K. Choi, Soriano, Zhang, & Rodríguez, 2012), while statins or fenofibrate are recommended in patients with hyperlipidaemia as they have a modest uricosuric effect (Derosa, Maffioli, & Sahebkar, 2015).

Prophylactic treatment is suggested once ULT has started. Usually, it consists of low doses of colchicine or low dose NSAIDs for a period of six months (P Richette et al., 2017).

Figure 1.5 summarises the pharmaceutical management's workflow in gout.

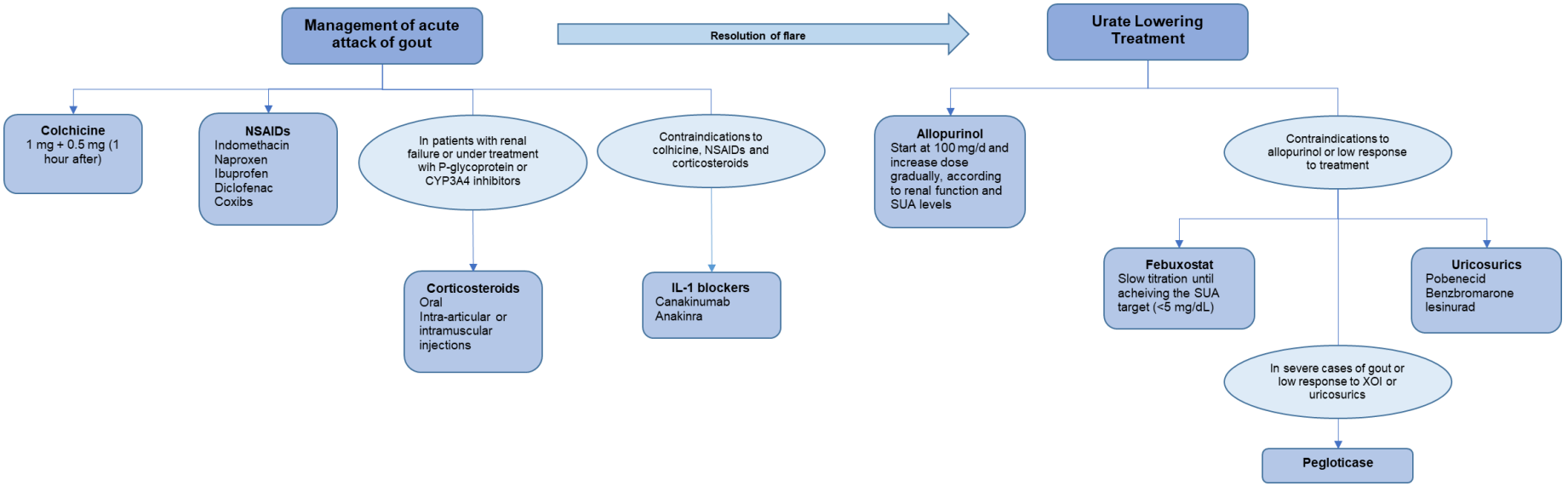


Figure 1.5 Management of gout.

Summary of pharmacological interventions according to the European league against Rheumatism and the American College of Rheumatology. NSAIDs: Nonsteroidal anti-inflammatory drugs. SUA: Serum uric acid. XOI: Xanthine oxidase inhibitors. Adapted from (Loeb, 1972; P Richette et al., 2017).

1.6 Pathogenesis of gout

1.6.1 Pathogenesis of hyperuricaemia

Purines are a group of molecules essential for multiple biochemical pathways, e.g. DNA and RNA synthesis. They are synthesised by both exogenous (*de novo* synthesis) and endogenous (salvage pathways) metabolism, interconverted into molecules that serve as intermediates in different pathways, and finally excreted as uric acid (UA) or allantoin (Figure 1.6), depending the species. (Jinnah, Sabina, & Van Den Berghe, 2013).

In humans, UA is the final product of purine metabolism because of the absence of the enzyme uricase. Uricase catalyses the oxidation from UA to allantoin, which is a more soluble metabolite. During primate evolution, the enzyme was lost due to two nonsense mutations that resulted in the truncation of the protein (Wu, D., Lee, & Caskey, 1992). Although the absence of uricase in humans is associated to protective effects - as UA prevents oxidative damage by scavenging free radicals (Oda, Satta, Takenaka, & Takahata, 2002); high levels of SU are also the main risk factor for other pathologies.

Under physiological conditions, UA can be found circulating in blood as soluble urate provided it does not exceed the saturation limit [6.8 mg/dL] (Nicola Dalbeth et al., 2016). To keep urate concentrations at soluble levels, it is crucial to maintain a sustained balance between purine intake, endogenous metabolism and excretion. When that balance is altered, UA concentration increases which can lead to urate precipitation, MSU crystals deposition within joints, and the eventual onset of gout. (Martillo et al., 2014)

Although it is well known that only around 10-12% of people with hyperuricaemia develop gout (Langford et al., 1987; B. J. Paul, Anoopkumar, & Krishnan, 2017), a high concentration of SU is still the main precursor to MSU

crystallisation (Nicola Dalbeth et al., 2016). Therefore, knowing the mechanisms behind hyperuricaemia is essential for the understanding of pathogenesis of gout.

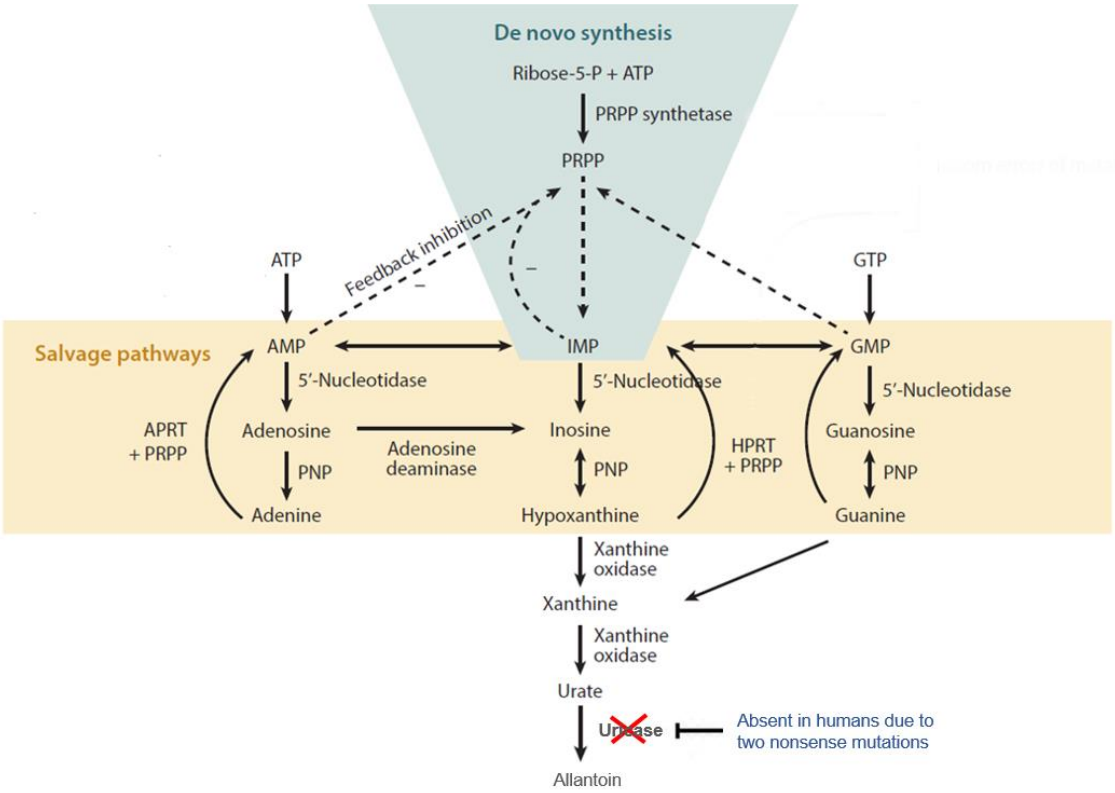


Figure 1.6 Purine metabolism pathway.

Purine synthesis starts with the presence of other precursors from the pentose phosphate pathway –ribose 5-phosphate- and involves both *de novo* and salvage metabolism. In the first case, the final product is inosine monophosphate (IMP), which is used to synthesise adenosine monophosphate (AMP) and guanosine monophosphate (GMP) via the salvage pathway. The final product that results from the degradation of purines is allantoin; however, in humans is uric acid due to the loss of the enzyme uricase. Adapted from (Mandal & Mount, 2015)

1.6.2 Role of hyperuricaemia in gout

Hyperuricemia can be caused by increased production or under-excretion of UA. Previously, it was classified as urate *overproduction hyperuricemia* and *renal underexcretion hyperuricemia*. The first type considered an increased production of UA due to a rich purine diet, or to other rare genetic causes. The

second type resulted because of a reduced renal excretion of UA; however, only two thirds of UA are eliminated by the kidney. The remaining is excreted by the intestine, which can also be compromised and result in hyperuricemia. (T. R. Merriman & Dalbeth, 2011)

With a better and more complete understanding of urate metabolism, hyperuricaemia is currently classified as renal underexcretion and renal overload, which groups urate overproduction and extra-renal underexcretion (Figure 1.7). Clinically, the type of hyperuricemia is defined according to the urinary urate excretion (UUE) values, and urate clearance/creatinine clearance ratio (also known as FE_{UA}). (Ichida et al., 2012; H. Matsuo et al., 2016)

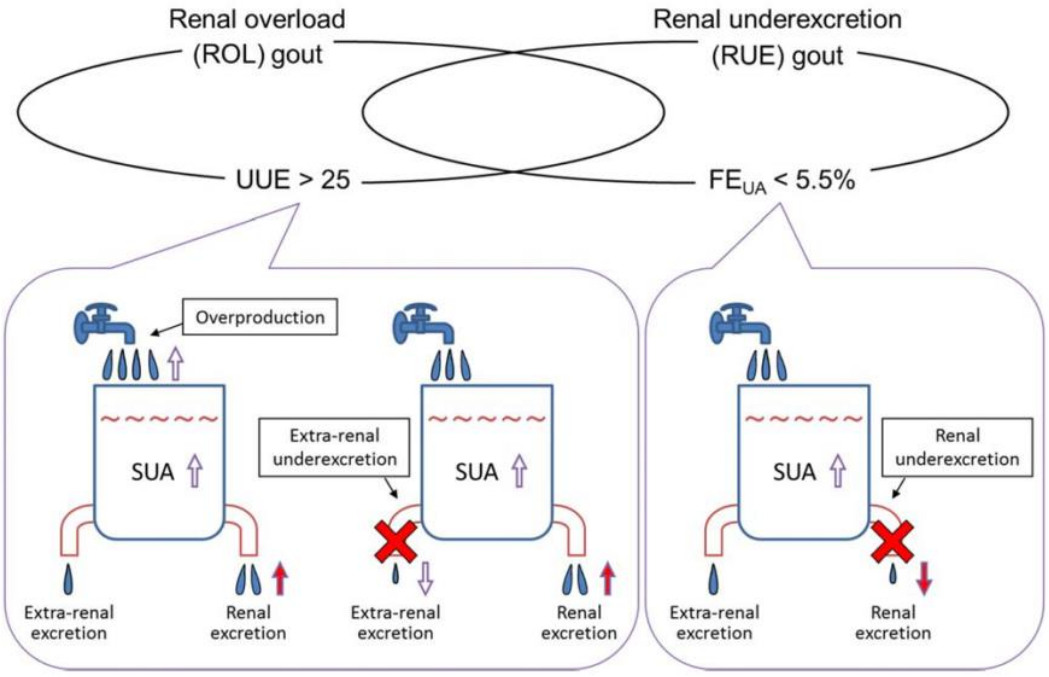


Figure 1.7 Classification of hyperuricemia.

Hyperuricemia can be classified as renal overload and renal underexcretion. Renal overload is caused by urate overproduction or extra-renal underexcretion, in both cases the urinary urate excretion (UUE) is >25 mg h^{-1} / 1.73 m^2 and the urate clearance/creatinine clearance ratio (FE_{UA}) is >5.5%. Renal underexcretion is caused by an impaired renal excretion, therefore it shows UUE <25 mg h^{-1} / 1.73 m^2 and FE_{UA} <5.5%. Taken from (Nakayama et al., 2017)

1.6.2.1 Renal overload hyperuricemia

Renal overload hyperuricemia could be caused by either overproduction of UA or reduced urate intestinal excretion. Renal overload hyperuricemia shows UUE values greater than 25 mg h^{-1} / 1.73 m^2, and FE_{UA} values greater than 5.5% (Ichida et al., 2012).

Besides diet with a high content of purines and fructose, urate overproduction can also be caused by failures in the endogenous metabolism of purines. There are several genetic disorders that affect the function of critical enzymes involved in the purine pathway (Figure 1.6).

Lesch-Nyhan syndrome is an X-linked disorder that affects the enzyme hypoxanthine-guanine phosphoribosyltransferase (HPRT). This enzyme is responsible of converting hypoxanthine and guanine into inosine monophosphate and guanosine respectively. Mutations in the *HPRT* gene result in a deficient enzyme and the subsequent accumulation of hypoxanthine and guanine, which are converted into UA. Milder consequences of the disease include hyperuricemia, gout and nephrolithiasis; while motor and intellectual disabilities are characteristic of a severe phenotype. (Jinnah et al., 2013; Torres & Puig, 2007)

Another X-linked syndrome is the phosphoribosyl pyrophosphate (PRPP) synthetase superactivity. PRPP synthetase uses ribose-5-posphate and ATP to catalyse the synthesis of PRPP. Superactivity of this enzyme is caused by mutations or overexpression of the gene *PRPS1* (J Roessler et al., 1994). As PRPP is produced at higher rates, it triggers an increased synthesis of UA. The clinical phenotype includes early onset hyperuricemia, hyperuricosuria (that if left untreated can evolve to gout and renal impairment), hearing loss, ataxia and hypotonia (Nyhan, 2005).

Although urate homeostasis can be compromised because of an overproduction of uric acid, this type of hyperuricemia represents only 10% of the cases, while the remaining results from urate underexcretion (H. K. Choi et al., 2005). The mechanisms of urate transport are complex, but it is well known that two-thirds of excretion occurs from in the kidneys and one third of uric acid is excreted via gastrointestinal tract. This occurs from urate transporters present in this tissue: these are important in the development of hyperuricemia and gout.

ABCG2 has been identified as a key UA transporter, located in the liver, the apical membrane of the intestine, and the proximal tubule (Woodward et al., 2009). Multiple GWAS have reported the single nucleotide polymorphism (SNP) rs2231142 in ABCG2 gene as a susceptibility locus for hyperuricemia and gout (C.-J. Chen et al., 2018; Dehghan et al., 2008; Kottgen et al., 2013; Hirotaka Matsuo et al., 2014; Wen et al., 2015). Rs2231142 is a missense mutation resulting in a substitution from glutamine to lysine at position 141 (Q141K) of ABCG2 gene (C.-J. Chen et al., 2018). This SNP has been associated to an increment of urinary urate output due to a decreased function (up to 53%) of ABCG2 in the excretion of urate via the intestine (Ichida et al., 2012). Rs2231142 is responsible for an increase of serum urate by 0.217 mg/dL (Kottgen et al., 2013).

1.6.2.2 Renal underexcretion hyperuricemia

Renal underexcretion hyperuricemia can be caused by either acquired factors or genetic variants that predispose to this disorder. It is characterised by UUE values less than 25 mg h^{-1} / 1.73 m^2, and FE_{UA} values less than 5.5% (Ichida et al., 2012).

The development of other metabolic disorders, chronic kidney disease, or the use of certain medications (diuretics) are among the acquired factors that can lead to a reduced glomerular filtration and hence to renal urate underexcretion (T. R. Merriman & Dalbeth, 2011).

On the other hand, genetic variants that contribute to a higher risk of renal underexcretion hyperuricemia are related to renal urate transport, which can be divided into four stages: (1) glomerular filtration, (2) reabsorption from the glomeruli, (3) secretion, and (4) postsecretory reabsorption in the proximal tubule (Mandal & Mount, 2015). Therefore, proteins involved in any of these stages have been subject to study. Moreover, GWAS have identified multiple variants located in urate transporter coding genes associated with renal underexcretion hyperuricemia.

SLC2A9 is the gene responsible for encoding the urate transporter GLUT9. It is found mainly in the proximal renal tubule and has been identified, along with ABCG2, as the genetic variant with most effect in gout and hyperuricemia. Although the exact causal variant has not been fully elucidated, it has been suggested that polymorphisms might be present at regulatory sites of SLC2A9 gene, affecting the expression of two different isoforms, which results in an increased urate reabsorption and decreased fractional excretion (Mandal & Mount, 2015). The SNP rs12498742, located in an intronic region of SLC2A9 gene has been pointed out as responsible for an increase in SU of 0.373 mg/dL (Kottgen et al., 2013). This variant, along with rs2231142 (ABCG2) contributes to a 3.4% of the genetic risk for hyperuricemia.

URAT1, another urate transporter, is encoded by SLC22A12 gene and is located in the proximal tubule of the renal cortex (Enomoto et al., 2002). URAT1 was one of the first urate transporters to be described, and thanks to its

importance in urate homeostasis, it has become an important drug target for several uricosurics such as probenecid and benzbromarone (Ragab et al., 2017)

1.6.3 Progression from hyperuricemia to gout

It is estimated that the heritability of hyperuricemia ranges from 40% to 70%; however, the genetic variants identified to date only explain 7% of the variation of uric acid levels (Dong et al., 2018). Moreover, though being crucial to gout, hyperuricemia does not always evolve to symptomatic gout. A follow-up study of patients with high SU levels (7.0–7.9 mg/dL) reported that during 14 years, only 12% presented gout.(Langford et al., 1987). Another study stated an annual incidence of gout of 0.1%, 0.5% and 4.9% in individuals followed up for 15 years, with SU levels of <7.0 mg/dL, 7.0-8.9 mg/dL and >9.0 mg/dL respectively (Campion, Glynn, & DeLabry, 1987).

Imaging methods have allowed the comparison of joint features in hyperuricaemic and normouricaemic individuals, looking for subclinical musculoskeletal involvement. There is evidence of low volumes of MSU crystal deposits in joints and tendons of people with hyperuricemia but without any signs of inflammation (Pineda et al., 2011). For instance, a DECT study showed a prevalence of MSU crystal deposition of 24% in patients with hyperuricemia (N. Dalbeth et al., 2015). Additional studies that used ultrasound and arthrocentesis to confirm presence of MSU crystals, reported a prevalence of 24.2% and 34.6%, respectively (Abhishek Abhishek et al., 2018; De Miguel et al., 2012). Asymptomatic hyperuricemia is therefore characterised by MSU crystal deposition but no clinical symptoms of gout. However, the mechanisms that lead from asymptomatic hyperuricemia to gout remain poorly understood.

Although there are several genome wide association analyses (GWAS) for gout (C. Li et al., 2015; H. Matsuo et al., 2016; J. Wang et al., 2012), they have been conducted in relatively small datasets, and most of them have identified risk variants that overlap with those associated to SU variation. Hitherto, there is only one GWAS of asymptomatic hyperuricaemia controls vs. gout cases in Japanese population (Kawamura et al., 2019). This study revealed two novel loci and a suggestive locus (rs7927466 in CNTN5, rs9952962 in MIR302F and rs12980365 in ZNF724) associated with the progression from asymptomatic hyperuricaemia to gout. The first two genes have been associated with other inflammatory diseases, which supports the importance of the immune response to urate crystals in the development of gout. Efforts have focused on the study of pathways such as the inflammasomes and Toll-like receptors, which are involved in the pathogenesis of the disease, to identify more genetic features that may contribute to the understanding of the progress from asymptomatic hyperuricemia to gout.

1.7 Immune mechanisms of MSU crystal induced inflammation

Inflammation is a consequence of an immune response against endogenous and exogenous factors that stimulate the activation of several signalling pathways and the secretion of pro-inflammatory molecules as a protective measure to avoid damage.

In gout, the release of MSU crystals from connective tissue and joint deposits is believed to be the inflammatory triggering factor, which induces the activation of both the innate and adaptive immune responses. Together, they cause the characteristic abrupt episodes of severe pain, swelling, redness and warmth in joints (R. Terkeltaub, 2017). Other factors such as changes in protein coating

on the crystals, changes in pro-inflammatory lipid coating etc. may also operate given the reported association between dietary and lifestyle excesses and gout flares (Ortiz-Bravo, Sieck, & Ralph Schumacher Jr, 1993; R. A. Terkeltaub, Dyer, Martin, & Curtiss, 1991), although these are not well understood.

Though the mechanisms by which MSU crystals induce inflammation are still under research, there is currently a better understanding of the pathways involved in the activation of the innate immune response, while the process in which the adaptive immunity participates in gout pathogenesis is less clear.

1.7.1 Innate immunity in gout

This type of immunity forms during embryonic development and evolves throughout time. It works as the first line defence from pathogens, either by physical barriers or by receptors that recognise pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMPs). Pattern recognition receptors (PRRs) can directly activate immune cells or promote the recruitment of macrophages, neutrophils, mast cells, etc. Subsequently, they induce the secretion of cytokines and expression of transcription factors to further trigger other immune responses, such as the complement and adaptive responses.(Hanson & Wigzell, 2014; Male, Brostoff, Roth, & Roitt, 2012)

MSU crystals in joints are identified as danger molecules by Toll-like receptors (TLRs) and nucleotide oligomerisation domain (NOD)-like receptors (NLRs). Both are responsible of the early immune response in gout. (N. Dalbeth & Haskard, 2005)

1.7.1.1 Toll-like Receptors

TLRs are a group of membrane glycoproteins that contain an extracellular region of leucine-reach repeat (LRR) motifs and an intracellular region known as Toll/IL-1R (TIR) domain. The first region is in charge of recognising a wide variety of ligands (e.g. lipopolysaccharides, envelope proteins, DNA fragments, RNA fragments, etc.), while the TIR domain's main function is to recruit the adaptor molecule myeloid differentiation primary response protein 88 (MyD88). MyD88 is a signal transducer that activates MAP kinases and allows translocation of the transcription factor NF-κβ to the nucleus to induce the expression of pro-inflammatory cytokines. This can be accomplished through a series of steps that include the recruitment of IRAK1 and IRAK4 and the consequent assembly of the complex TRAF6-TAB2-TAK1-TAB1 (Figure 1.8). (Akira & Takeda, 2004)

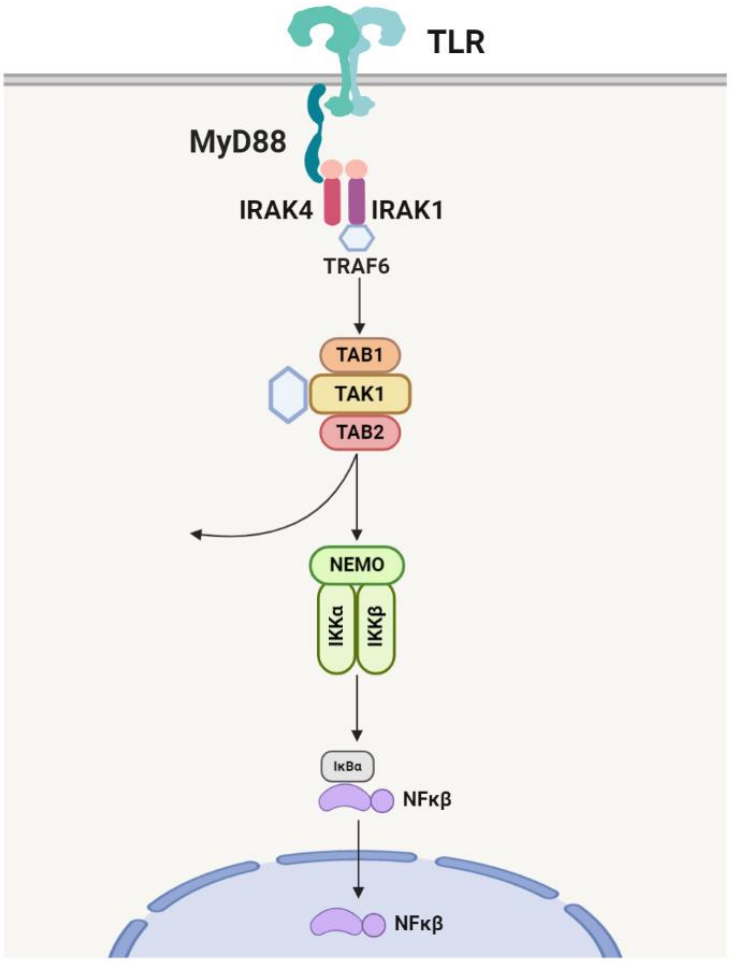


Figure 1.8 Toll-like receptors (TLRs) signalling pathway.

TLRs are a type of pattern recognition receptor that mediates the recruitment of MyD88. It triggers the association of IRAK4, which further phosphorylates IRAK1 and recruits TRAF6. After a series of phosphorylations and ubiquitylations, the complex TRAF6-TAB2-TAK1-TAB1 activates MAP kinases and IKK complex, which allows the translocation of transcription factor NF-κβ to the nucleus. (Akira & Takeda, 2004)

1.7.1.2 NOD-like receptors

On the other hand, NLRs are a different type of PRRs that form a tripartite complex of proteins, known as inflammasomes, that mediate the synthesis of caspases, which in turn cleave the pro-inflammatory cytokines IL-1β and IL-18 to produce their active forms (Latz, Xiao, & Stutz, 2013). Up to date, 22 NLRs have been identified (Man Si & Kanneganti, 2015).

In general, NLR are integrated by three domains: (1) a C-terminal region consisting of LRR motifs, (2) a nucleotide binding or NACHT domain, (3) and an amino-terminal domain that could be a caspase activation and recruitment domain (CARD), a pyrin domain (PYD), or a baculovirus inhibitor of apoptosis domain (BIR) (Menu & Vince, 2011). Under the presence of DAMPs, the LRR domain activates NLRs inducing a conformational change that allows the recruitment of the apoptosis speck protein (ASC). ASC functions as an adaptor protein, since most NLRs contain PYD instead of CARD, which is essential for NLR-pro-caspase-1 interactions. In this way, ASC's CARD domain binds to its homologous domain of pro-caspase-1. This oligomerisation prompts self-cleavage of pro-caspase-1 into p20-caspase-1 and p10-caspase-1 fragments, with the subsequent processing of pro-IL-1β and pro-IL-18 cytokines (Latz et al., 2013; Rathinam, Vanaja, & Fitzgerald, 2012). Figure 1.9 shows NLRP3 structure and its general signalling pathway, as an example of an ASC-dependent NLR.

The inflammasomes are one of the innate immune complexes that has been widely studied owing their role in several inflammatory, autoimmune and metabolic disorders (Guo, Callaway, & Ting, 2015). Out of the existing NLR inflammasomes, NLRP3 is best characterised, as it recognises an extensive number of PAMPs (e.g. *Stapylococcus aureus, Candida albicans, Influenza A* ligands) and endogenous DAMPs (e.g. cholesterol, calcium pyrophosphate dihydrate, hyaluronan, MSU crystals) (Jo, Kim, Shin, & Sasakawa, 2016). Furthermore, NLRP3 has been associated with the immune response of numerous pathologies such as diabetes, obesity, cancer, Parkinson, Alzheimer's disease, rheumatoid arthritis, and gout. (Menu & Vince, 2011).

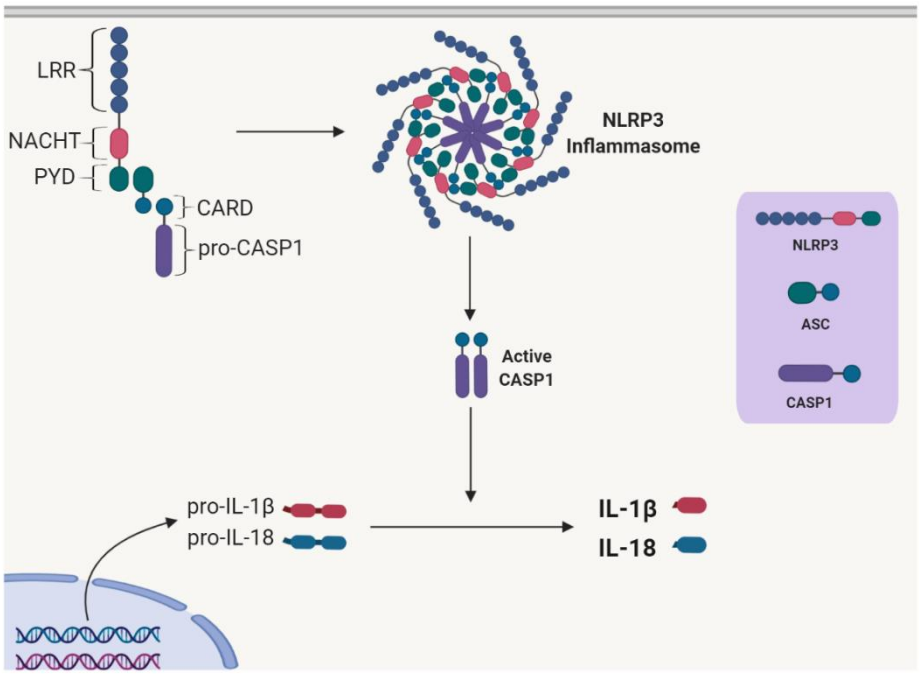


Figure 1.9 NLRP3 inflammasomes structure and signalling.

Most inflammasomes consist of three main molecules (NLR, ASC and pro-caspase-1) that facilitates the activation of inflammatory cytokines. NLRP3 inflammasome has three domains in its NLR component: LRR, NACHT and PYD. After activation, NLR recruits ASC, which contains a PYD domain and a CARD domain which finally interacts with pro-caspase-1. This multimeric protein assembly activates pro-caspase-1 through its self-cleavage into p10-caspase-1 and p20-caspase-1, which process the pro-inflammatory cytokines IL-1β and IL-18. (Rathinam et al., 2012)

In gout, TLR2, TLR4 and NLRP3 are the pattern recognition receptors that participate in the downstream signalling to activate IL-1β and IL-18 (Liu-Bryan, Scott, Sydlaske, Rose David, & Terkeltaub, 2005; Sutterwala, Haasken, & Cassel, 2014). Although they are critical in gout, PRRs cannot activate the immune response by themselves, but they need several biochemical changes, that together allow the oligomerisation of the inflammasome and the activation of transcription factors to finally promote the synthesis of pro-inflammatory cytokines (Kingsbury, Conaghan, & McDermott, 2011). These changes will be discussed with more detail in section 1.7.3.

1.7.2 Adaptive immunity in gout

Adaptive immunity is also known as acquired immunity, and unlike the innate response that only recognises highly conserved structures, it recognises antigens with high specificity. It is divided in two different groups according to the type of lymphocytes, as cellular (T cells) and humoral immunity (B cells) (Male et al., 2012). Furthermore, once activated they can undergo a clonal selection, which generates the immunologic memory (Scherer & Burmester, 2011).

T cells develop in the thymus and respond to intracellular pathogens. They are divided into two main types of cells: CD4+ T-cells and CD8+ T-cells. CD4+ T-cells consist of five types including the T helper cells Th_1, Th_2, Th_{17}, follicular helper cell, and the regulatory treg cells. Depending on the type, they can directly interact with phagocytes (Th_1) and B-cells (Th_2) to underpin antigen presentation or activate the synthesis of pro-inflammatory cytokines such as IL-17 (Th_{17}) to recruit other immune cells. CD8+ T cells are divided into two types: the cytotoxic T-cells and CD8+ T_{REG} cells. The former promotes the destruction of host cells under the presence of intracellular pathogens, while the second limit the immune response to prevent overreaction against self-tissues. (Male et al., 2012; Overgaard Nana, Jung, Steptoe Raymond, & Wells James, 2014)

B cells on the other hand, respond to extracellular pathogens. However, to initiate their response they require other signals mediated by T-cells or dendritic cells. Depending on the type of signal, they can further differentiate into memory cells or plasma cells. The latest are in charge of producing IgM and IgD antibodies, which successively evolve to the subclasses IgA, IgE and IgG. These have higher affinity that confers precise functions to achieve a specific immune response. (Bonilla & Oettgen, 2010)

According to several studies, the adaptive immunity plays a role in different stages of gout pathogenesis. First, it has been observed that humoral immunity might participate in the crystallisation of MSU and the subsequent activation of the inflammatory response (Martillo et al., 2014). Terkeltaub et al described that after incubating MSU crystals with plasma and serum, among the proteins that were bound to the crystals were IgGs and complement-associated proteins (R. Terkeltaub, Tenner Andrea, Kozin, & Ginsberg Mark, 1983). Subsequent studies not only confirmed the existence of specific antibodies against MSU, but also correlated the concentration of immunoglobulins with urate crystallisation rates (Kaneko & Maru, 2000; Kanevets, Sharma, Dresser, & Shi, 2009). However, these results are based on *in vitro* observations, thus the humoral response *in vivo* remains hypothetical.

Other markers of the adaptive immunity that have been associated with the inflammatory mechanisms in gout are the Th17 and CD8+ T cells. Th17 cells promote the synthesis of IL-17, which has been found in synovial fluid of patients with gout (Kotake et al., 1999). A study conducted by Conforti et al showed that MSU crystals activate dendritic cells that in turn stimulate Th17 differentiation (Conforti-Andreoni et al., 2011). Interestingly, this differentiation only occurs in the presence of IL-1, which is dependent on NLRP3 inflammasome assembly and the expression of NF-κβ, suggesting a central role of IL-1 in MSU crystal induced inflammation (Chung et al., 2009).

Although the role of CD8+ T-cells is less clear, since MSU crystals recruit dendritic cells, it has been suggested that they may activate T-cells (Ghaemi-Oskouie & Shi, 2011). In addition, it has been found that tophus contain abundant CD8+ cells with the RANKL (receptor activator of nuclear factor κβ ligand) receptor expressed in their membranes. RANKL is responsible for the

activation of oesteoclasts. Thus, CD8 cells might have a role in advanced stages of gout by prompting bone erosion in patients with tophi (S.-J. Lee et al., 2011).

Even though several markers of the adaptive immunity have been identified in gout, suggesting an active response of both humoral and cellular systems, their contribution to disease pathogenesis remains to be fully elucidated. Yet, these studies have highly contributed to building a clearer picture of the pathways involved in MSU crystals-induced inflammation.

1.7.3 Mechanism of inflammation in gout

The hallmark immune sign in gout is the synthesis of pro-inflammatory cytokines upon exposure of monocytes (macrophages) and surrounding synovial cells to urate crystals. These cells excrete interleukins IL-1β and IL-18, which amplify the inflammatory response by attracting neutrophils to the synovium (Kingsbury et al., 2011). The activation of the immune response involves the collective action of multiple pathways that activate different signals at both the transcriptional and post-translational level.

The activation of the NLRP3 inflammasome requires two essential steps (Figure 1.10). The first one consists of an increase in the expression of NLRP3 protein, mediated by TLR-MyD88 pathway (activated once the LRR domain of TLRs interacts with MSU crystals). The second step comprises the assembly of the inflammasome components NLRP3, ASC and pro-caspase-1. It has been identified that after macrophages phagocyte DAMPs, the assembly and subsequent activation of the inflammasome complex depends on three different signals: (1) presence of reactive oxygen species (ROS), (2) decrease in

potassium (K+) concentrations, and (3) synthesis of cathepsin B. (Jo et al., 2016)

Once MSU crystals are ingested, macrophages enter a stage of cellular stress characterised by an increase of ROS. These potentiate the dissociation of the antioxidant TRX (thioredoxin) and its inhibitor TXNIP (TRX-interacting protein) (Hwang et al., 2014). As soon as TXNIP is released, it interacts with NLRP3 inflammasome priming its activation (Zhou, Tardivel, Thorens, Choi, & Tschopp, 2009). In addition, intracellular urate crystals induce mitochondrial impairment, resulting in an abundant excretion of ROS that directly stimulate NLRP3 inflammasome (Gurung, Lukens, & Kanneganti, 2015).

Under normal conditions, macrophages present high concentrations of intracellular K+. However, it has been observed that DAMPs phagocytosis changes cell osmolality, decreasing potassium concentrations significantly. This might be in part due to an increased production of ATP by impaired mitochondria. ATP molecules activate the potassium membrane transporters P2X7 and Pannexin-1, which are responsible for potassium efflux from the cell. (Petrilli et al., 2007; Sutterwala et al., 2014)

Finally, the third signal is mediated by lysosomal damage. Once ingested, MSU crystals accumulate inside lysosomes, causing their membrane disruption and excretion of enzymes with proteolytic activity, such as cathepsin B. This enzyme not only promotes cell death, but also serves as an additional signal for inflammasome activation. (Lima Jr et al., 2013)

Although the exact mechanisms by which these intracellular changes interact or stimulate NLRP3 activation remain unclear, it has been observed that they create an optimal environment for its assembly (Tschopp & Schroder, 2010).

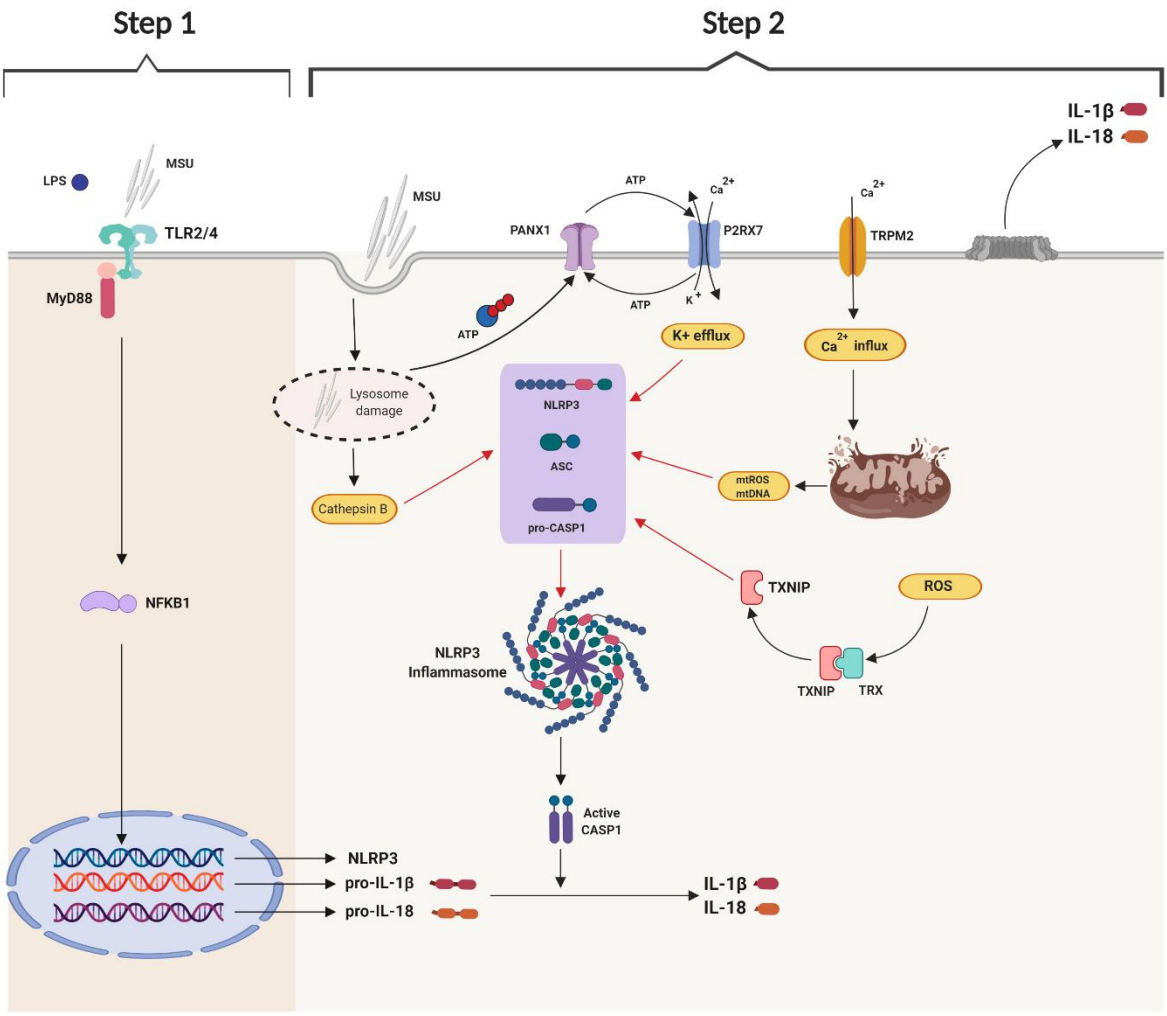


Figure 1.10 Assembly and activation of the inflammasome NLRP3. Adapted from (Jo et al., 2016)

While NLRP3 inflammasome is activated, extracellular MSU crystals prompt TLR2 and TLR4 activities. As mentioned in section 1.4.1, DAMPs interact with the LRR domain of TLRs, initiating the recruitment of MyD88. This pathway induces the expression of NF-κβ, which upregulates the transcription of IL-1β and IL-18. Initially, these interleukins are excreted from the nucleus as their non-activated forms (pro-IL-1β and pro-IL-18), but to be excreted from the cell, they need to be cleaved to their active forms. This is the stage when the NLRP3 inflammasome is crucial, since it is in charge of synthesising caspase-1. Caspase-1 is the enzyme that cleaves pro-IL-1β and pro-IL-18 to the active

cytokines IL-1β and IL-18. (N. Dalbeth & Haskard, 2005; Ghaemi-Oskouie & Shi, 2011; Kingsbury et al., 2011)

Once pro-inflammatory cytokines are excreted from macrophages, they interact with other receptors (e.g. IL-1R) from surrounding cells of the synovium. These receptors, just as TLRs, when activated recruit MyD88, upregulating the expression of transcription factors and the synthesis of more cytokines and chemokines (Kingsbury et al., 2011; Xiao et al., 2015). This process amplifies the inflammatory response through the influx of neutrophils to the site of MSU crystals deposition and the activation of the adaptive response. Figure 1.11 shows the mechanisms involved in gouty inflammation.

On the other hand, once neutrophils ingest urate crystals, they form extracellular aggregates (also known as neutrophils extracellular traps (NETs)) that eventually lead to the resolution of the inflammatory response. NETs are highly compacted structures of engulfed crystals surrounded by a group of molecules (histones, DNA, anti-inflammatory proteins and enzymes) that dissolve MSU crystals and inactivate the synthesis of pro-inflammatory cytokines. (Schett et al., 2015; Schorn et al., 2012)

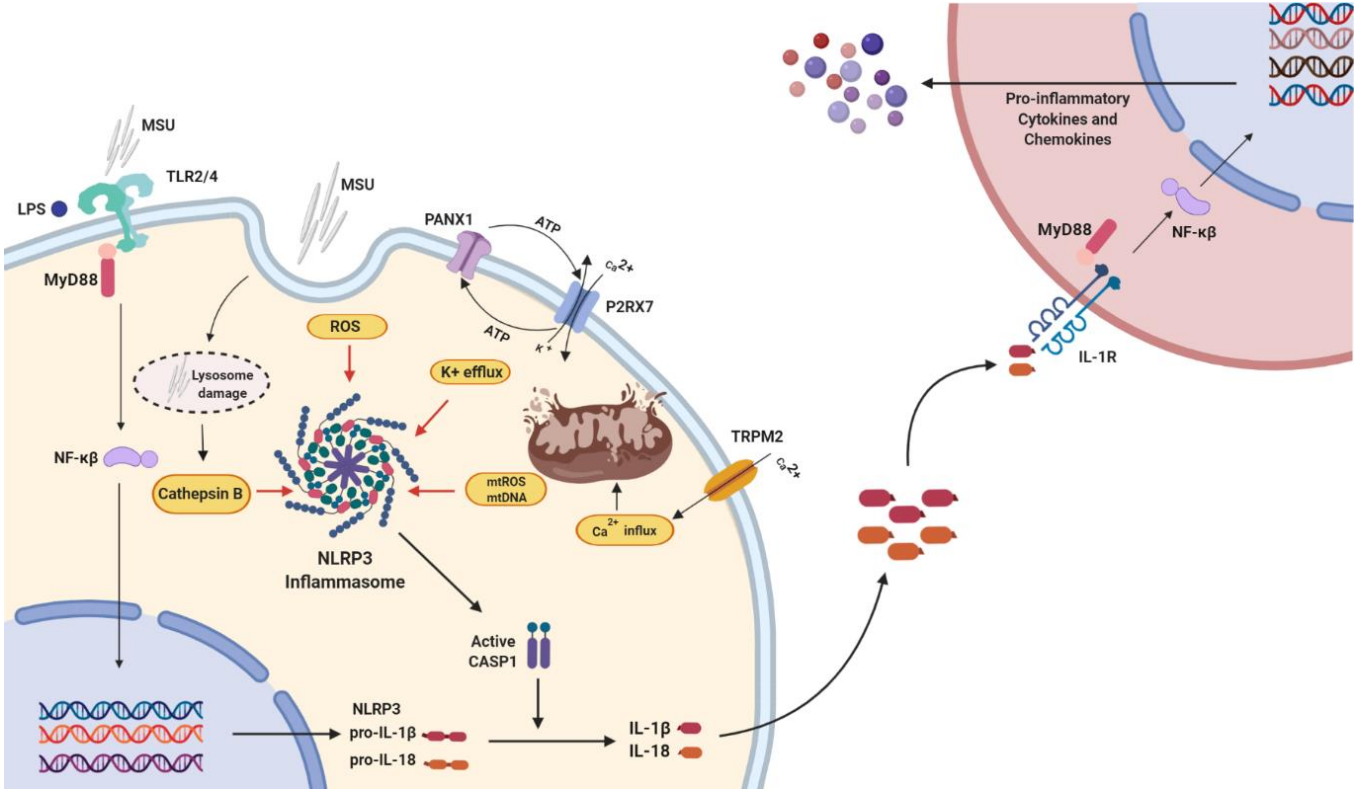


Figure 1.11 Molecular mechanisms of inflammation in gout.
Adapted from (Kingsbury et al., 2011)

1.8 Genetic studies in hyperuricaemia and gout

The understanding of the molecular mechanisms in gout has been facilitated by genetic studies that have identified relevant pathways and loci involved in the development of the disease. Among these studies are candidate gene association studies and genome wide association studies (GWAS), which have been essential to the research of genetics in complex diseases (Jeck, Siebold, & Sharpless, 2012). GWAS permits the investigation of genotype information of thousands of common variants (minor allele frequency –MAF >5%) across the entire genome and their association with a particular phenotype (Manolio, 2010).

A GWAS published in 2007 (S. Li et al., 2007), was the first to associate polymorphisms in SLC2A9 gene ($P=1.84x10^{-16}$) with hyperuricemia. Interestingly, SLC2A9 had not been previously related to urate metabolism; however, subsequent GWAS confirmed its association with SU levels ($P=2x10^{-15}$, OR=1.89 [1.36,2.61]) (C. Wallace et al., 2008) and its role in urate transport (Vitart et al., 2008).

Further GWAS also identified the polymorphisms rs2231142 (P = 9.0x10-20, β = 0.25 [0.03]) and rs1165205 (P = 5.6x10-10, β = -0.11 [0.02]), in ABCG2 and SLC17A3 genes respectively, to be associated to high levels of UA. This study also explored SNPs associations with gout, but only rs2231142 reached statistical significance (Dehghan et al., 2008).

The first GWAS meta-analysis, comprised of 28,141 individuals, reported seven additional signals that influence SU concentrations: SLC17A1 ($P=3.04x10^{-14}$, β=-0.06 [-0.078,0.459]), SLC16A9 ($P=1.07x10^{-8}$, β=0.078 [0.051,0.105]), SLC22A11 ($P=6.68x10^{-14}$, β=0.062 [0.046,0.078]), SLC22A12 ($P=2.04x10^{-9}$, β=0.056 [0.074,0.038]), PDZK1 ($P=2.68x10^{-9}$, β=-0.062 [-0.083,-0.042]), GCKR

(P=1.4x10^{-9}, β=0.052 [0.035,0.068]) and LRRC16A (P=8.5x10^{-9}, β=0.054 [0.036,0.072]) (Kolz et al., 2009). These signals were later confirmed in other populations (Y. H. Lee & Song, 2012; Nakayama et al., 2013; A. Tin et al., 2011; J. Wang et al., 2012).

The CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology) consortium found two new SNPs in INHBC (P=1.9x10^{-11}, β=5.16 [0.77]) and RREB1 (P=1.0x10^{-9}, β=4.39 [0.72]) (Yang et al., 2010). These loci were replicated in a meta-analysis of 48 GWAS of serum urate levels and 14 GWAS of gout, comprising 110,347 and 2,115 Caucasian individuals respectively (within the Global Urate Genetics Consortium) (Kottgen et al., 2013). This meta-analysis revealed 18 novel susceptibility loci in addition to the 12 already reported. Table 1.1 summarises association results for the novel variants for urate concentrations, gout results are not included as none of those novel loci showed genome-wide significance.

Other GWAS have successfully identified four additional loci associated with gout in Japanese and Chinese populations: ALDH2 (P=1.7x10^{-18}, OR=0.53 [0.37,0.52]) (Masayuki Sakiyama, Matsuo, Nakaoka, et al., 2016), BCAS3 (P=1.36x10^{-13}, OR=0.79), RFX3 (P=1.48x10^{-10}, OR=0.81) and KCNQ1 (P=1.28x10^{-8}, OR=0.82) (C. Li et al., 2015).

In 2019, Tin *et al* published a trans-ancestry meta-analysis of GWAS for SU variation in a total of 457,690 individuals. This meta-analysis revealed 183 risk loci, of which 147 were novel. These loci were also tested for association with gout in a meta-analysis of 20 GWAS (n=13,179 gout cases). Fifty five variants had a significant association (P<2.7x10^{-4}), with ABCG2 showing the greatest OR (2.04 [1.96-2.12]). Additionally, the 183 SU index SNPs were used to predict gout risk in an independent dataset using the UK Biobank (n=334,880 in total

and 4,908 gout cases), adding age and sex into the demographic model. The genetic model alone had an area under the receiving operator characteristic (AUROC) curve of 67%, compared to 80% for the demographics model, and84% for the combined model. This study also analysed genetic correlation of SU levels with more than 700 complex diseases, showing that besides gout (r=0.92), several cardiometabolic traits, such as insulin resistance, dyslipidaemia and obesity, presented the highest correlation coefficients. Finally, Tin et al also reported that HNF4A regulates ABCG2 transcription in the kidney and that the functional variant rs1800961 in HNF4A lowers serum urate variation via an increased urate excretion mediated by ABCG2. (Adrienne Tin et al., 2019)

Although GWAS have contributed to the identification of new genetic variants associated to both hyperuricemia and gout, only a minority of those SNPs have shown a clear biological implication in the development of the diseases. A limitation of GWAS is that the effect of the identified variants tends to be small, leaving much of the missing heritability unexplained. Moreover, due to linkage disequilibrium, GWAS findings cannot be certainly interpreted as causal (Manolio, 2010).

Given that most genetic variants identified to date are located in regulatory regions, expression analysis of GWAS discovery sets might contribute to elucidate true causative variants and their functional effects (Tony R. Merriman, 2015).

Table 1.1 Serum urate GWAS meta-analysis.
Novel susceptibility loci reported by (Kottgen et al., 2013).

Gene	Variant	P-value	Beta	SE
TRIM46	rs11264341	6.2×10^{-19}	-0.05	0.006
INHBB	rs17050272	$1.6X10^{-10}$	0.035	0.006
SFMBT1	rs6770152	$2.6X10^{-16}$	-0.044	0.005
TMEM171	rs17632159	$3.5X10^{-11}$	-0.039	0.006
VEGFA	rs729761	$8.0X10^{-16}$	-0.047	0.006
BAZ1B	rs1178977	$1.2X10^{-12}$	0.047	0.007
PRKAG2	rs10480300	$4.1X10^{-9}$	0.035	0.006
STC1	rs17786744	$1.4X10^{-8}$	-0.029	0.005
HNF4G	rs2941484	$4.4X10^{-17}$	0.044	0.005
A1CF	rs10821905	$7.4X10^{-17}$	0.057	0.007
ATXN2	rs653178	$7.2X10^{-12}$	-0.035	0.005
UBE2Q2	rs1394125	$2.5X10^{-13}$	0.043	0.006
IGF1R	rs6598541	$4.8x10^{-15}$	0.043	0.006
NFAT5	rs7193778	$8.2x10^{-10}$	-0.046	0.008
MAF	rs7188445	$1.6x10^{-9}$	-0.032	0.005
HLF	rs7224610	$5.4x10^{-17}$	-0.042	0.005
ACVR1B-ACVRL1	rs7976059	$1.9x10^{-9}$	0.032	0.005
B3GNT4	rs7953704	$2.6x10^{-8}$	-0.029	0.005

1.9 Rationale of the study

Despite genetic studies that have contributed to understand the molecular mechanisms involved in the development of hyperuricemia and gout, the reasons underlying the progress from asymptomatic hyperuricemia with or without MSU crystal deposition to symptomatic gout remain unclear.

This project aims to investigate whether the expression of genes involved in different inflammatory pathways known to be activated during gout flares (e.g. inflammasome and TLR) also associate with the transition from asymptomatic hyperuricemia with/without crystal deposition to gout.

Additionally, to build on the work done by other investigators and following on from the release of serum urate data from UK Biobank, this study will for the first-time attempt to ascertain the genetic factors that influence the transition from hyperuricaemia to gout. This will be done using GWAS and polygenic risk score (PRS) approaches. This we anticipate will improve the understanding of why some people with hyperuricemia develop symptomatic gout while others do not and whether there is an influence of inflammatory gene expression or inflammatory gene variants in this transition.

1.10 Aims and objectives

The overall purpose of this project is to examine genetic factors associated with the transition from hyperuricaemia to gout.

The specific objectives are:

a) To evaluate whether a differential expression of inflammasome and TLR genes is associated with SU levels and transition from asymptomatic hyperuricaemia to gout.
b) To determine inflammasome and TLR gene expression profile during acute and intercritical gout.
c) To generate a PRS model to distinguish gout cases from controls regardless of serum urate.
d) To generate and validate a GWAS of asymptomatic hyperuricaemia vs gout.
e) To generate a PRS model to distinguish gout cases from asymptomatic hyperuricaemia controls.

The objectives will be met in two distinct studies. The first one involves wet-lab work to explore gene expression profiles and cytokine measurements in patients at different stages of gout pathogenesis, recruited at the Department of Academic Rheumatology (Chapter 2). The second study involves dry-lab work using phenotype and genotype data from the UK Biobank resource to generate a PRS model of gout vs. controls (Chapter 3), and to conduct association studies in gout cases vs. asymptomatic hyperuricaemia controls (Chapter 4).

Chapter 2. Gene and cytokine profiling in individuals with normouricaemia, asymptomatic hyperuricaemia and gout.

2.1 Introduction

Hyperuricaemia is the main precursor to gout. Even though environmental factors such as diet and lifestyle influence both serum urate level and gout risk, there is a genetic component influencing their development with an estimated heritability of 40-70% (Nicola Dalbeth et al., 2016; Nath et al., 2007). Several groups worldwide have conducted genetic studies to investigate the genes contributing to the development of hyperuricaemia and gout. For nearly a decade, genome wide association studies (GWAS) have been the key methodology used (Kawamura et al., 2019; Kottgen et al., 2013; C. Li et al., 2015; H. Matsuo et al., 2016; Adrienne Tin et al., 2019; Voruganti et al., 2013). However, thus far the genes identified as risk loci, explain only 6-7% of serum urate (SU) variation, because genetic variants tend to exert only small effects that are not easily captured with GWAS (Dong et al., 2018). On the other hand, functional studies have also investigated in detail other processes associated to gout, based on the knowledge of the mechanisms involved in different stages of its development (C. J. Chen et al., 2006; Holzinger et al., 2014; Kanevets et al., 2009; Liu-Bryan et al., 2005; Martin William, Walton, & Harper, 2008; McKinney et al., 2015).

The progression of gout can be divided in four stages, as summarised in figure 2.1. The first one is hyperuricaemia, traditionally defined as SU levels >6.8 mg/dL (408 µmol/L) without evidence of monosodium urate (MSU) crystals deposits. Followed by the deposition of MSU crystals within joints without symptoms of gout (i.e. without any gout flares). The third is the deposition of

MSU crystals with recurrent episodes of acute inflammation (gout flares). The last stage is advanced gout, characterised by chronic arthritis, joint damage, and presence of tophi. (Nicola Dalbeth & Stamp, 2014)

The factors governing the transition from hyperuricaemia to asymptomatic MSU crystal deposition, and from that state to symptomatic gout are poorly understood. This is despite the fact that the causative mechanisms of hyperuricemia are well understood with several genetic variants that contribute to renal overload or renal under excretion of urate identified (T. R. Merriman & Dalbeth, 2011). It is also known that during a gout flare, the immune response to MSU crystals is specifically mediated by TLRs and the NLRP3 inflammasome (Liu-Bryan, 2010). Whether variations in these genes and their expression affects the transition from asymptomatic hyperuricaemia to gout is not known.

Thanks to the improved knowledge of the immune mechanisms prompted by MSU crystals as a crucial step in the development of gout flare, several studies have explored them into more detail as a mean to address the existing gaps. Various groups have analysed the impact after MSU crystals stimuli on the expression of genes and proteins such as *CASP-1, CCL2, CXCL1, IL-1β, IL-6, IL-8, IL-37, NLRP3, SOCS3, TNF-α*, etc., that are involved in the inflammatory response. However, most of them have been conducted either *in vitro*, by extracting synoviocytes that are subsequently exposed to MSU crystals (Dang, Xu, Xie, & Zhou, 2018; Anna Scanu et al., 2010; Y. Yu et al., 2019), or *in vivo* using murine models of gout (L. Liu et al., 2016; A. Scanu et al., 2015; Y. Yu et al., 2019). Although, some other studies have investigated gene expression and cytokine profiles in synovial fluid or serum of controls and cases with acute and intercritical gout. (Alberts et al., 2019; Holzinger et al., 2014; Kienhorst et al., 2015; Yang Liu et al., 2018), only two have examined cytokines and leukocyte

counts in individuals with asymptomatic deposition of MSU crystals. These studies reported that levels of IL-6 and IL-8 were higher in patients with asymptomatic hyperuricaemia (mean (SD) = 69.7 (73.5) pg/mL and 18.5 (25.7) pg/mL, respectively), compared to normouricaemic controls (mean (SD) = 28.2 (17.6) pg/mL and 7.5 (6.1) pg/mL, respectively). Additionally, leukocyte counts were higher in patients with asymptomatic deposition of MSU crystals (median (IQR) = 200 (138-540) mm^3) compared to controls (median (IQR) = 30 (10-53) mm^3) (Andrés, Bernal, Arenas, & Pascual, 2019; Estevez-Garcia et al., 2018).

Therefore, the aim of the current study was to analyse mRNA and cytokine concentration in different stages of pathogenesis of gout, i.e. normal serum urate without MSU crystal deposition, hyperuricaemia without MSU crystal deposition, hyperuricaemia with asymptomatic MSU crystal deposition, inter-critical gout and gout flares.

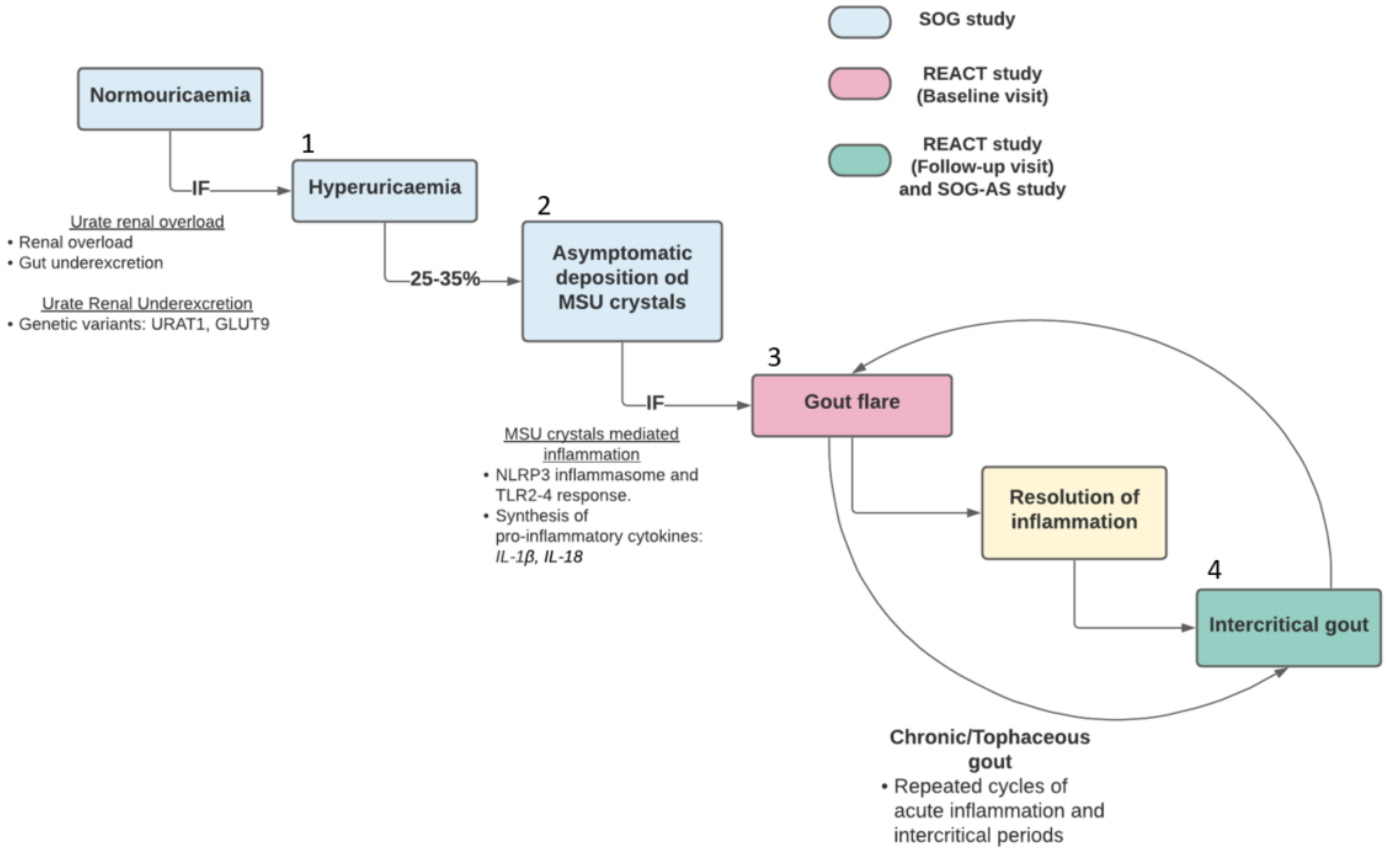


Figure 2.1 Stages in the progression of gout.

2.2 Materials and Methods

2.2.1 Sample source

This research involved the analysis of gene expression and cytokine measurement for different sample sets, corresponding to three clinical studies conducted at the Department of Academic Rheumatology, Nottingham City Hospital, University of Nottingham. These studies were subjected to ethical approval from the University of Nottingham medical school or the NHS Research Ethics Committees. References 15/SC/0730, 15/EM/0316 and 18/EM/0324.

2.2.1.1 Sons of gout study (SOG)

This study involved men whose father or mother had gout (Abhishek Abhishek et al., 2018). For recruitment, people with gout who participated in previous clinical research studies (and agreed to be contacted for future studies) at Nottingham University or attended the Nottingham NHS Treatment Centre were mailed with a study pack to post to their sons. Additionally, the study was advertised on Facebook and in one local newspaper.

In total 1,435 people with gout were sent a study invitation pack to pass to their son(s). Interested individuals (249 sons >20 years of age) either returned a reply slip or contacted the research team in response to advertisements. They were asked to complete a screening questionnaire providing relevant clinical data to exclude those who could be classified as having gout based on the Mexico criteria (Peláez-Ballestas et al., 2010). Of the 249 who replied, 134 were contactable or willing to attend a research visit, and of these, 131 met the entry criteria and were considered eligible for a study visit. Demographic information, lifestyle data and clinical history (including comorbidities and drug prescriptions)

were recorded. Blood and urine samples were also collected during the study visit.

As the objective of the sons of gout study was to estimate the prevalence of asymptomatic MSU crystal deposition, participants were assessed by musculoskeletal ultrasonography of target sites (blind to SU levels) to look for double contour sign, intra-articular tophi or aggregates in the joints and tendons (Figure 2.2).

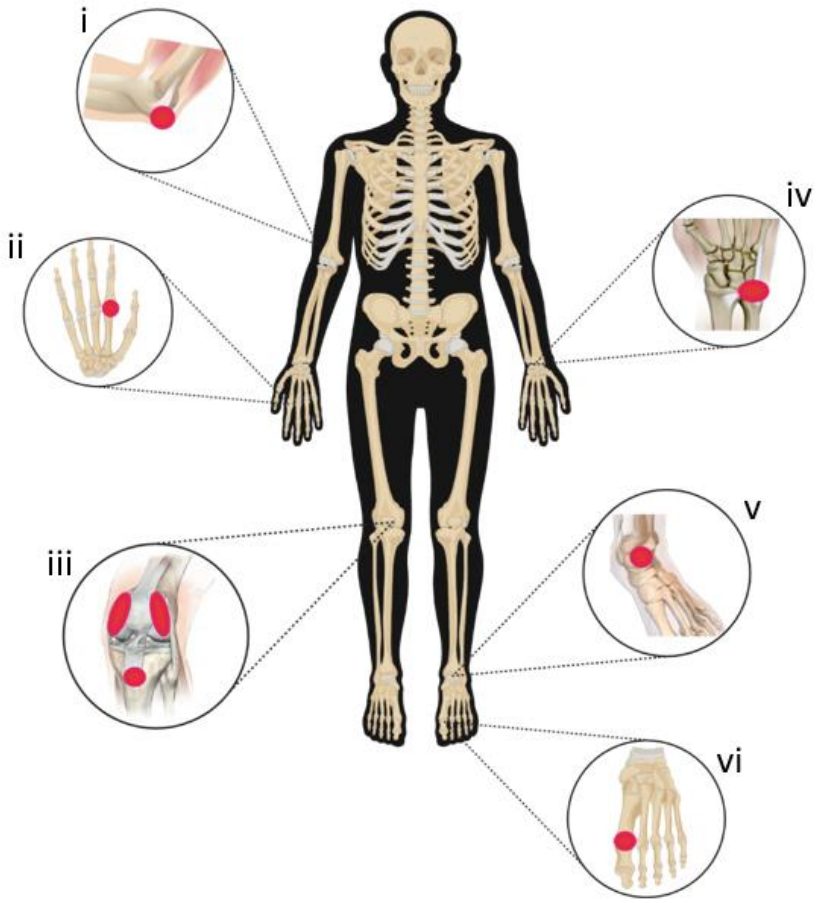


Figure 2.2 Ultrasonographic assessment.

The following joints were assessed by ultrasound: (i) triceps tendon insertion, (ii) second metacarpophalangeal joints, (iii) femoral condyles, (iv) wrist triangular fibrocartilages, (v) talar domes, and (vi) first metatarsophalangeal joints. (Abhishek Abhishek et al., 2018)

2.2.1.2 Immune response in acute gout: REACT study

This study comprised patients with gout (diagnosis confirmed by arthrocentesis, ultrasound, plain radiography showing erosive tophaceous gout or clinical assessment by a rheumatologist for tophi) who presented with a gout flare of no more than seven days' duration before the baseline visit. Subjects were recruited by three different means: (1) From among inpatients with gout flare admitted to Nottingham University Hospitals NHS Trust. (2) Patients with gout seen at the Nottingham University Hospitals NHS Trust or at the Circle Nottingham NHS Treatment Centre were requested to contact the research team when they developed a gout flare. (3) The study was also advertised with posters in doctor's offices in Department of Rheumatology Nottingham University Hospitals NHS Trust.

Patients willing to participate were contacted by a member of the research team, who screened them for eligibility. Twenty-one patients met the entry criteria and underwent a baseline visit, during which demographic data, clinical history and gout-specific information were recorded. Joint assessment, blood and urine sample collection were also performed. Out of the 21 patients recruited, 19 (90.5%) agreed to give samples. Serum samples were used to measure cytokine levels, and three additional tubes (serum separator, DNA isolation and RNA stabiliser) were collected for further analyses.

Participants were followed up at weeks 6 and 12 after the baseline visit, where they were asked to update information about flares' symptoms and/or treatment, and provide blood and urine samples.

2.2.1.3 Switch off gout attack study (SOG-AS)

This randomised placebo controlled study was designed to determine whether Omega-3 fatty acids prevent gout flares in patients starting on urate lowering treatment (ULT). Participants were recruited by three different means: (1) Via primary care, where GP surgeries acted as participant identification sites. (2) Via secondary care, where hospitals acted as participant identification sites. (3) The study was also advertised with posters in clinics, pharmacies, newspapers or on social media.

Individuals willing to participate were invited to a screening visit. Briefly, participants were eligible if they: (1) met the American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria for gout (Neogi et al., 2015), (2) were willing to commence ULT, (3) had SU levels ≥360 µmol/L, and (4) presented at least one flare in the previous twelve months. Demographic data, clinical history and gout related data were collected from participants who met the entry criteria. In total, 56 participants agreed to provide blood samples for cytokines measurements, and for genetic studies (optional). Participants were randomly prescribed an Omega-3 or a placebo, they were asked to start on allopurinol or febuxostat until the 5th week. Dose up-titration visits were arranged every 2-3 weeks until the participants reached the treatment target of SU ≤300 µmol/L. Number of flares, changes in medication and side effects related information was collected during all visits; additionally, blood samples for cytokines, and for genetic studies (if agreed by the participant) were also collected. Only data from the baseline samples are included in this study.

2.2.2 Sample collection

Blood samples for RNA extraction were collected in PAXgene® Blood RNA Tubes according to manufacturer's instructions. These tubes contain an RNA stabilizer reagent to improve RNA quality and quantity. Samples were stored at -80°C in Academic Rheumatology Unit facilities and transferred later to Human Genetics Laboratory facilities (Life Sciences Building, University Park) for further use. The samples were transferred on dry ice.

Samples for cytokine measurements were collected in serum separator tubes. These were centrifuged and serum was aliquoted into 2.0 mL tubes (two per sample). Serum samples were stored at -80°C in Academic Rheumatology Unit facilities and shipped later to Affinity Biomarker Labs facilities for cytokine measurements.

2.2.3 RNA extraction

Total RNA was purified using QIAGEN's PAXgene® Blood RNA kit following the manual RNA extraction protocol. This protocol initiated with an incubation step of PAXgene® Blood RNA tubes at room temperature for two hours. Tubes were then centrifuged for 15 minutes at 5000 x g to collect leukocytes pellets, which were subsequently resuspended in 4.0 mL RNase-free water, and centrifuged again for 15 minutes at 5000 x g. The pellet was resuspended in 350 µL BR1 buffer and transferred into a 1.5 mL microcentrifuge tube; 300 µL of BR2 buffer and 40 µL proteinase K were added to allow the lysis of cellular membranes and digestion of proteins respectively. This sample was mixed and incubated for 10 minutes at 55°C using a thermoshaker at 1400 rpm. The lysate was transferred to a spin column and centrifuged for 3 minutes at 18000 x g. The supernatant was then transferred to a 1.5 mL microcentrifuge tube without

disturbing the pellet, and 350 µL ethanol were added in order to optimise nucleic acids' binding properties. This solution was passed through a selective silica membrane, to which nucleic acids bind, by centrifuging for 1 minute at 18000 x g. After a wash step with BR3 buffer, the membrane was treated with 80 µL DNase I incubation mix (10 µL DNase I + 70 µL DNA digestion buffer) for 15 minutes at room temperature. The spin column was then washed with 350 µL BR3 buffer by centrifuging for 1 minute at 18000 x g. Two additional wash steps using 500 µL BR4 buffer were conducted, centrifuging for 1 minute and 3 minutes during each step at 18000 x g. Finally, RNA was eluted in a final volume of 40 µL BR5 elution buffer, and incubated at 65°C for 5 minutes. A 5.0 µL aliquot was taken from each RNA sample for further quality control analysis. All RNA samples were stored at -80°C.

2.2.4 RNA quality control

All RNA samples were quantified on the Agilent BioAnalyser 2200 TapeStation. RNA integrity and DV_{200} values were also examined. For this protocol, 1.0 µL of RNA (or RNA ScreenTape ladder) mixed with 5.0 µL ScreenTape buffer were used. Samples were heated at 72°C for 3 minutes and placed on ice for 2 minutes. After centrifuging, samples were loaded into the BioAnalyser with an RNA ScreenTape. The TapeStation consists of a simplified electrophoresis that outputs an electropherogram and gel-like figure, which provides a RIN^e score (RNA integrity number equivalent) as a quantitative measure of RNA quality. Samples are required to have a RIN^e score ≥6.0 to be considered acceptable for cDNA synthesis. Furthermore, DV_{200} values were also calculated from the 2200 TapeStation controller. DV_{200} values give a more sensitive measure of RNA quality for gene expression analyses, by considering the mean RNA

fragment size. Values represent the percentage of RNA fragments greater than 200 nucleotides.

Samples that did not have a RIN^e score greater than 6.0, or a DV_{200} value above 70% were not considered for cDNA synthesis.

2.2.5 cDNA synthesis

Complementary DNA (cDNA) synthesis was performed using QIAGEN RT^2 First Strand Kit. Before the reverse transcription reaction, this protocol included an additional step to remove residual genomic DNA (gDNA). If primers for gene expression analyses are not designed across exon-exon junctions, they amplify gDNA, therefore its elimination is crucial to avoid false positive results.

gDNA elimination started with 500 ng (set as the starting amount for all samples) of RNA, mixed with 2.0 µL reaction buffer GE and RNase-free water (if necessary) to get a final volume of 10 µL. The reaction mix was incubated at 42°C for 5 minutes and placed on ice for at least 1 minute. The total volume (10 µL) of gDNA elimination mix was then added to the reverse transcription mix, which consisted of 4.0 µL 5x BC3 buffer, 3.0 µL RNase-free water, 2.0 µL RE3 reverse transcriptase mix, and 1.0 µL external P2 control (a built-in external control that can be detected by an assay in the QC PCR array to test for possible inhibition during the reverse transcription reaction). The final 20 µL mix was then incubated for 15 minutes at 42°C, and 5 minutes at 95°C. Finally, 91 µL RNase-free water were added to each reaction, mixed gently and stored at -80°C. A 7.0 µL aliquot was taken from each cDNA sample for further viability analysis.

2.2.6 cDNA viability

Before gene expression analysis, RNA and cDNA quality was verified using the QIAGEN RT^2 RNA QC PCR arrays. These plates allow the simultaneous analysis of 12 samples to assess the following parameters:

RNA integrity: The arrays include primers to amplify two differentially expressed housekeeping genes (ACTB and HPRT1). This gives an idea of the threshold cycle (C_T) ranges expected for the expression array.

gDNA contamination: two assays are included to evaluate this parameter. One uses total RNA as starting material to amplify a housekeeping gene (GDC), while the other includes a set of primers specifically designed to amplify non-transcribed DNA (NRT).

Reverse transcription inhibition (RCT): to test for possible reverse transcription inhibitors generated during RNA purification, one well is designed to amplify the external control (P2 control) supplied on the RT^2 first strand kit.

PCR inhibition (PPC): the array contains two wells (per sample) with a plasmid template control and primers to amplify it. cDNA is added into one of the wells, and RNase-free water into the remaining well.

DNA contamination: A non-template control (NTC) is also included to detect the introduction of any DNA contamination during the array setup.

The experimental setup involved preparation of three PCR component mixes, which were transferred into the plate according to the manufacturer's layout (Figure 2.3). Briefly, the first mix included 6.0 µL cDNA, 75 µL 2x RT^2 SYBR Green mastermix, and 69 µL RNase-free water. A volume of 25 µL mix 1 was dispensed to the first five rows of the array (A-E). The second mix added 1.0 µL total RNA (diluted 1:100), 13 µL 2x RT^2 SYBR Green mastermix, and 13 µL

RNase-free water. 25 µL mix 2 were transferred to row F. The third mix contained 30 µL 2x RT^2 SYBR Green mastermix, and 30 µL RNase-free water. 25 µL mix 3 were transferred to rows G and H as shown in figure 2.3. Finally, the array was sealed with 8-cap strips, mixed for 1 minute using a 96-well plate shaker, and centrifuged for 1 minute using a microplate centrifuge.

As RNA quality was carefully confirmed on the BioAnalyser Tape Station, only 48 samples were randomly selected to be analysed on the RT^2 RNA QC PCR arrays. Amplification was performed in the Agilent AriaMx Thermocycler, under the following conditions: 1 cycle for 10 minutes at 95°C, 40 cycles: 15 seconds at 95°C, 1 minute at 60°C (+data collection), and a final dissociation cycle (30 seconds at 95°C, 30 seconds at 60°C, and 30 seconds at 95°C with a resolution of 0.5 °C/sec).

Sample		1	2	3	4	5	6	7	8	9	10	11	12
cDNA Template	A	ACTB	ACTB	ACTB	ACTB	ACTB	ACTB	ACTB	ACTB	ACTB	ACTB	ACTB	ACTB
	B	HPRT1	HPRT2	HPRT3	HPRT4	HPRT5	HPRT6	HPRT7	HPRT8	HPRT9	HPRT10	HPRT11	HPRT12
	C	RTC	RTC	RTC	RTC	RTC	RTC	RTC	RTC	RTC	RTC	RTC	RTC
	D	PPC	PPC	PPC	PPC	PPC	PPC	PPC	PPC	PPC	PPC	PPC	PPC
	E	GDC	GDC	GDC	GDC	GDC	GDC	GDC	GDC	GDC	GDC	GDC	GDC
RNA	F	NRT	NRT	NRT	NRT	NRT	NRT	NRT	NRT	NRT	NRT	NRT	NRT
RNase-free water	G	PPC	PPC	PPC	PPC	PPC	PPC	PPC	PPC	PPC	PPC	PPC	PPC
	H	NTC	NTC	NTC	NTC	NTC	NTC	NTC	NTC	NTC	NTC	NTC	NTC

Figure 2.3 QIAGEN RT^2 RNA QC PCR array layout.

Threshold cycles (C_T) were calculated for each well by defining a threshold above the background signal but in the lower half of the linear phase. This threshold was consistent across all RT^2 RNA QC PCR arrays. Additionally, PCR specificity was analysed through the inspection of melting curves for each well,

which must have shown a single peak at temperatures greater than 80ºC. C_T values were exported to a spreadsheet, and those showing values greater than 35 were reported as negative calls. QC parameters were analysed as follows:

Housekeeping gene expression: Both ACTB and HPRT1 housekeeping genes included in the array not only provided an estimate of C_T values expected for the expression array, but also guarantied RNA integrity. Therefore, C_T values lower than 35 were expected for all samples.

Reverse transcription inhibition: If inhibitors were present during the reverse transcription reaction, they would affect the amplification of the external control into the cDNA synthesis mix. Therefore, a ΔC_T was calculated by subtracting the sample C_T minus the control C_T ($\Delta C_T = C_T^{RTC} - C_T^{PPC\ (RNase-free\ water)}$). A ΔC_T greater than 5 was indicative of apparent inhibition.

PCR inhibition: Similarly to the previous control, if impurities were present during the PCR, the sample C_T would be different from the RNase-free water well C_T. Therefore, a ΔC_T ($\Delta C_T = C_T^{PPC(cDNA)} - C_T^{PPC\ (RNase-free\ water)}$) greater than 3 indicated inhibition during the amplification reaction.

gDNA contamination: This parameter required observing C_T^{NRT} and C_T^{GDC}, if the first one showed a value ≥ 35, then no gDNA contamination was present. However, if the value was less than 35, but C_T^{GDC} was greater than 35, the remaining gDNA would not affect the gene expression results.

General contamination: To guarantee no general DNA contamination was introduced during the plate setup, C_T^{NTC} values should have been ≥ 35.

2.2.7 Gene expression profiling

Gene expression was analysed using a customised 96-well Human RT^2 Profiler PCR Array. Each plate is comprised of 86 inflammasome and TLRs associated genes, 3 housekeeping genes (ACTB, B2M, and RPLP0), 1 control to test gDNA contamination (GDC), 3 controls to assess reverse transcription inhibition (RTC), and 3 positive PCR controls (PPC) (Figure 2.4). This array is based on the QIAGEN Human Inflammasomes RT^2 Profiler PCR Array. However, its content was modified to include the following genes: *CARD8, CD14, DAPK1, MAP3K11, NEK7, TLR2*, and *TLR4*. These genes were selected because they either play key roles in both the regulation of NLRP3 inflammasome and the immune mechanisms in gout, or they have been identified in gout GWAS and candidate gene association studies (Y. He, Zeng, Yang, Motro, & Nunez, 2016; H. Matsuo et al., 2016; McKinney et al., 2015; Razmara et al., 2002).

This customised array was designed to run only one sample per plate. Therefore, only one PCR mix was needed for each plate. The mix included 102 µL cDNA, 1350 µL 2x RT^2 SYBR Green mastermix, and 1248 µL RNase-free water. The PCR mix was transferred to a loading reservoir, and 25 µL were dispensed into each well of the array using an electronic multichannel pipette. The plate was sealed with 8-cap strips, mixed for 1 minute, and centrifuged for 1 minute. PCR was performed in the Agilent AriaMx thermocylcer under the following conditions: 1 cycle for 10 minutes at 95°C, 40 cycles: 15 seconds at 95°C, 1 minute at 60°C (+data collection), and a final dissociation cycle (30 seconds at 95°C, 30 seconds at 60°C, and 30 seconds at 95°C with a resolution of 0.5 °C/sec).

A threshold above the background signal but in the lower half of the linear phase was defined, making sure it remained consistent across all RT^2 Profiler PCR

arrays. Specificity for all wells was assessed by verifying a single peak appearing in the melting curves. C_T values were calculated for each well and exported to a spreadsheet. A C_T value of 35 was defined as the cut-off.

Quality control was thoroughly assessed per plate; this involved the calculation of the average C_T values for all the controls. A value less than 5 in the difference of the average C_T^{RTC} minus the average C_T^{PPC} indicated no reverse transcription inhibition. The average C_T^{PPC} should have not varied by more than 2 cycles to guarantee the absence of PCR inhibitors. Finally, C_T^{GDC} values greater than 35 indicated the absence of gDNA contamination. Only the samples that met these quality control criteria were considered for relative expression analyses.

	1	2	3	4	5	6	7	8	9	10	11	12
A	AIM2	BCL2	BCL2L1	BIRC2	BIRC3	CARD8	CASP1	CASP5	CASP8	CCL2	CCL5	CCL7
B	CD14	CD40LG	CFLAR	CHUK	CIITA	CTSB	CXCL1	CXCL2	DAPK1	FADD	HSP90AA1	HSP90AB1
C	HSP90B1	IFNB1	IFNG	IKBKB	IKBKG	IL12A	IL12B	IL18	IL1B	IL6	IRAK1	IRF1
D	IRF2	MAP3K7	MAP3K11	MAPK1	MAPK11	MAPK12	MAPK13	MAPK3	MAPK8	MAPK9	MEFV	MYD88
E	NAIP	NEK7	NFKB1	NFKBIA	NFKBIB	NLRC4	NLRC5	NLRP1	NLRP12	NLRP3	NLRP4	NLRP5
F	NLRP6	NLRX1	NOD1	NOD2	P2RX7	PANX1	PEA15	PSTPIP1	PTGS2	PYCARD	PYDC1	MOK
G	RELA	RIPK2	SUGT1	TAB1	TAB2	TIRAP	TLR2	TLR4	TNF	TNFSF14	TNFSF4	TRAF6
H	TXNIP	XIAP	ACTB	B2M	RPLP0	HGDC	RTC	RTC	RTC	PPC	PPC	PPC

Figure 2.4 QIAGEN Customised 96-well Human RT^2 Profiler PCR Array layout

2.2.8 Relative expression analysis

The $\Delta\Delta C_T$ method (Livak & Schmittgen, 2001) was used to calculate fold changes for each gene. Data was normalised to RPLP0 housekeeping gene, which showed less variation among all samples than the other two housekeeping genes.

First, the $\Delta\Delta C_T$ method involves the calculation of the difference between the C_T values of the gene of interest (GOI) and the C_T of the housekeeping gene (RPLP0).

$$\Delta C_T = C_T^{GOI} - C_T^{RPLP0}$$

Then, the difference in the average ΔC_T values of each gene (biological replicates) between groups is calculated.

$$\Delta\Delta C_T = \Delta C_T^{Group\ 2} - \Delta C_T^{Group\ 1}$$

(group 2 is the experimental set and group 1 is the control set)

Finally, fold changes are calculated as $2^{(-\Delta\Delta C_T)}$. A fold change greater than 1 is indicative of an upregulation, while values between 0 and 1 indicate downregulation.

2.2.9 Cytokine measurement

The cytokines that were selected for quantitation were: GRO-α, hsCRP, IL-1β, IL-18, IL-6, IL-8, IP-10, MCP-1, TNF-α, and VEGF-α. Cytokines were selected if their RNA levels showed a differential expression, or based on their importance in the MSU crystal induced inflammatory response.

A total of 244 serum samples were shipped to Affinity Biomarkers Labs facilities, where biomarker analyses were performed. Cytokines were measured using two

different platforms: Latex agglutination for hsCRP, and Meso Scale Discovery (MSD) U-Plex Array for GRO-α, IL-1β, IL-18, IL-6, IL-8, IP-10, MCP-1, TNF-α, and VEGF-α. A brief description of each method is outlined below.

Latex agglutination: this method was conducted using the CardioPhase® hsCRP kit, and analysed on the ADVIA Clinical Chemistry System. This semi-quantitative method employed latex particles coated with anti-human CRP antibody, that when mixed with serum containing CRP, agglutinated and caused changes in the suspension's turbidity. This turbidity was measured at 571 nm, and compared to the standard curve values to determine CRP concentration.

MSD U-Plex Array: the principle of this assay is similar to the sandwich ELISA, but it allows multiplex analyses by using biotinylated capture antibodies (specific to each biomarker), which are coupled to U-plex linkers. These linkers self-assembled onto exclusive sites on the plate, and serum samples were then added. After the antigen bound to the capture antibody, a second antibody (conjugated with electrochemiluminiscent labels) was added. The plate was finally analysed in an MSD instrument, which applies voltage to the plate electrodes, causing light emission from the labels. Each biomarker's concentration was measured based on the light intensity.

2.2.10 Statistical analysis

Samples from SOG and REACT studies were used for both gene expression analysis and cytokine measurements. Additionally, samples from SOG-AS study were also used for cytokine measurements. Since the number of samples for RNA extraction and serum collection were different for all studies, due to either lack of

consent of the patients to perform genetic studies or failure of quality control parameters for gene expression profiling, the number of participants for gene expression and cytokine measurements varied. Therefore, the demographic data for each study are shown for each analysis in the corresponding results section.

The mean ± standard deviation (SD), and number (n) and proportion (%) were used to describe demographic and comorbidity data respectively. Body mass index (BMI) was calculated from weight and height, measured during baseline visit as weight/height^2 [kg/m^2]. One-way Analysis of Variance (ANOVA) and chi-square test were used to compare continuous ad categorical data, respectively.

All statistical analyses were performed using SPSS® Statistics version 24. Graphs were produced using GraphPad Prism version 8.

2.2.10.1 Gene expression Statistics

Relative expression (ΔC_T) for SOG study was compared among the following groups: (1) normal SU levels (<360 µmol/L) without MSU crystal deposits; (2) high SU (≥360 µmol/L) without MSU crystal deposits; and (3) high SU (≥360 µmol/L) with MSU crystal deposits. The Kruskal-Wallis H Test with Bonferroni correction for pairwise comparisons was used to determine statistical significance. Although hyperuricaemia is usually defined as SU >406 µmol/L (6.8 mg/dL), this is the concentration at which urate crystallises *in vitro* at physiological conditions. However, a threshold of 360 µmol/L (6.0 mg/dL) was chosen, given its association with incident gout (Nicola Dalbeth et al., 2018).

Relative expression between the baseline visit and the follow-up visit of REACT study, was compared using the Wilcoxon signed-rank test.

P values were corrected for multiple testing using a false discovery rate (FDR) of 5%. Adjusted p values (P_{FDR}) less than 0.05 were considered significant.

2.2.10.2 Cytokine measurement Statistics

Raw data for all cytokine measurements were verified to replace values marked as out of the detection limit to a value below the cut-off.

Cytokine levels were compared for REACT groups using the Friedman test, adjusted by Bonferroni correction for pairwise comparison. Additionally, cytokine levels for the three groups from SOG study, the active flare group from REACT study, and intercritical gout group from SOG-AS study were compared using the Kruskal-Wallis H test with Bonferroni correction for pairwise comparisons. P values were adjusted for multiple testing using a 5% FDR, and adjusted p values less than 0.05 were considered significant.

2.2.10.3 Sample size and power estimates

Calculating power estimates and sample sizes for gene expression analyses is challenging, due to the number of genes per array and number of groups under study. Wei *et al* propose a procedure to estimate power and sample size for gene expression studies following a gene-by-gene basis to determine the pooled standard deviation among groups, for pairwise comparisons (Wei, Li, & Bumgarner,

2004). The R function to determine the required sample size to detect a 2-fold change with an 80% power is:

```
power.t.test (n=x, delta=1, sd=0.81, sig.level = 0.001, power=0.8, type = "two.sample", alternative = "two.sided")
```

Where: delta=1 represents a 2-fold change in expression level between the groups under study and the control group; sd= the 75th percentile of the pooled standard deviation; sig.level= p value <0.001 to account for multiple testing; power= 80%. Solving for this function, with a p-value of 0.001 and 80% power, at least 26 samples per group are required to detect a 2-fold change in the 75% least variable genes.

2.3 Results

2.3.1 Gene expression profiling in individuals with normouricaemia, hyperuricaemia without MSU crystal deposition, and MSU crystal deposition without gout

2.3.1.1 SOG study: Sample processing and demographic characteristics

In total, 131 participants met the entry criteria and underwent a study visit. Of these, 113 gave consent for collecting blood samples for genetic analysis. Demographic characteristics and comorbidities of the 113 participants included in this study are summarised in Table 2.1. Participants were classified according to SU levels and presence of MSU crystals deposits. In total, 33 participants had normal SU levels (<360 µmol/L); 60 had high SU (≥360 µmol/L) without MSU crystal deposits; and 20 had high SU (≥360 µmol/L) and presence of MSU crystal deposits. The mean (SD) age, BMI, and SU were 45.7 (10.9) years, 27.2 (4.8) kg/m^2**,** and 391.2 (67.8) µmol/L respectively. The age and BMI were comparable among the three groups.

The prevalence of hypertension in this population was 9.7%, while diabetes and hyperlipidaemia had a prevalence of 1.8% and 6.2% respectively.

Total RNA was extracted, with all RNA samples passing the quality control parameters for cDNA synthesis. A total of 48 RNA and cDNA samples were randomly selected to assess their viability prior to gene expression profiling. All 48 samples passed the quality control parameters of the QIAGEN RT^2 RNA QC PCR arrays. However, after analysing the internal controls included in the customised Human RT^2 Profiler PCR Arrays, 21 samples did not meet either the reverse transcription or the PCR efficiency controls. Table 2.1 shows demographics of included/excluded cases.

Table 2.1 Demographic characteristics and comorbidities of the Sons of Gout Study participants.

	Total	**SU<360μmol/L**		**SU≥360μmol/L**		**SU≥360μmol/L + MSU crystal deposits**	
	n=113	Included (n=31)	Excluded* (n=2)	Included (n=44)	Excluded* (n=16)	Included (n=17)	Excluded* (n=3)
Age, years, mean (SD)	45.7 ± 10.9	48.1 ± 12.4	48.5 ± 2.1	44.3 ± 10.1	43.12 ± 12.2	47.2 ± 9.2	45.0 ± 12.8
BMI, kg/m^2, mean (SD)	27.2 ± 4.8	26.5 ± 6.1	26.2 ± 5.7	27.5 ± 4.1	28.1 ± 5.4	27.1 ± 3.3	26.8 ± 2.6
SU, μmol/L, mean (SD)	391.2 ± 67.8	312.9 ± 37.3	273.5 ± 78.5	429.4 ± 44.7	414.9 ± 36.9	428.4 ± 49.1	408.5 ± 27.2
Comorbidities, n (%)							
Hypertension	11 (9.7)	6 (19.4)	0	1 (2.3)	4 (25.0)	0	0
Hyperlipidaemia	7 (6.2)	3 (9.7)	0	2 (4.5)	1 (6.3)	1 (5.8)	0
Diabetes	2 (1.8)	1 (3.2)	0	0	1 (6.3)	0	0

SU: Serum urate; BMI: Body mass index; MSU: Monosodium urate. *Cases were excluded if they did not meet the reverse transcription or the PCR efficiency QC.

2.3.1.2 SOG study: Gene expression profiling

To determine whether there was a differential expression of inflammasome and TLRs associated genes among the three study groups, real-time RT-PCR was performed using a customised 96-well Human RT^2 Profiler PCR Array. A fluorescence threshold of 0.3 was defined for all arrays, and C_T values were calculated for each well. Amplification data was normalised to *RPLP0* housekeeping gene, and relative expression among groups was compared using the Kruskal-Wallis H test with Bonferroni post-hoc test. After correcting for multiple testing using a 5% FDR, 25 genes had an adjusted P value <0.05. However, six genes (*CCL2, CCL7, DAPK1, IFNG, IL12B, IL6, MAPK11, NLRP4*, and *NLRP5*) had a threshold cycle relatively high (>33.5) in all groups, meaning that the expression level in PMBCs was low. Therefore, for this study they were not considered significant since a greater sample size is recommended to validate results. From the remaining 19 genes, only 10 (*BIRC2, CD40LG, CXCL1, 1L-1β, MEFV, NLRP12, PANX1, TNFSF14, TXNIP*, and *XIAP*) had a fold change ≥1.5 for at least one of the hyperuricaemia groups compared to the normouricaemia group (Figure 2.5). IL-1β had the greatest fold changes for both SU ≥360 µmol/L and SU ≥360 µmol/L + MSU crystals groups, compared to the control group (2.27 and 3.79, respectively), followed by MEFV (1.79 and 3.25, respectively) and CXCL1 (1.46 and 2.90, respectively). Table 2.2 summarises the fold-changes for each gene and their Kruskal-Wallis H test p values.

Table 2.2 Fold changes in gene expression for SU ≥360 µmol/L and SU ≥360 µmol/L + MSU crystals, compared to SU <360 µmol/L

	SU ≥360 µmol/L	**SU ≥360 µmol/L + MSU crystals**	
Gene	**Fold-change**	**Fold-change**	**Adj. p value***
BIRC2	1.16	1.86	0.010
CD40LG	1.29	1.55	0.039
CXCL1	1.34	2.75	0.011
IL-1β	2.10	3.58	0.031
MEFV	1.67	3.10	0.011
NLRP12	1.27	1.82	0.036
PANX1	0.91	2.20	$2.753x10^{-05}$
TNFSF14	1.00	1.90	0.010
TXNIP	1.15	1.79	0.024
XIAP	1.00	1.79	0.001

Relative expression was compared using the Kruskal-Wallis H test; and correction for multiple testing was conducted using a 5% FDR. *5% FDR adjusted p value.

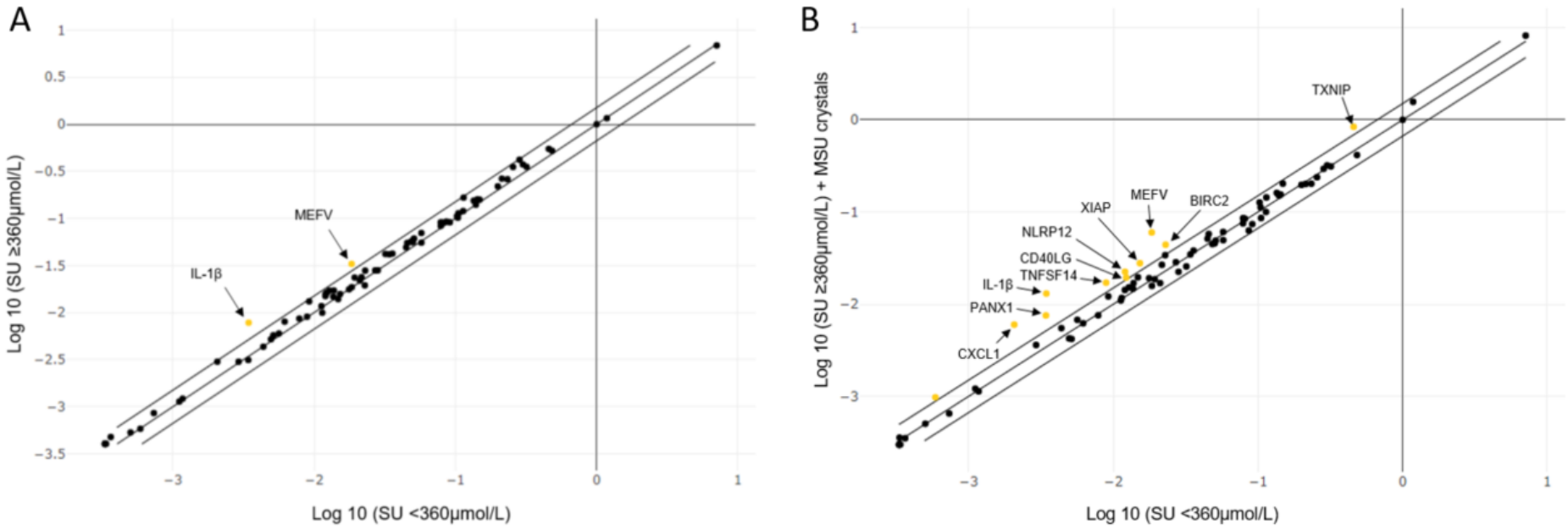


Figure 2.5 SOG: Scatter plot of the log base 10 ($2^{-\Delta CT}$) of each gene of the Human RT^2 Profiler PCR Array.

The middle line indicates a fold change of 1.0, and upper and bottom lines indicate up and down regulation respectively. Yellow dots represent upregulated genes with a fold change >1.5. A) SU ≥360 µmol/L vs SU <360 µmol/L. B) SU ≥360 µmol/L +MSU crystals vs SU <360 µmol/L.

Pairwise comparisons of the relative expression (ΔC_T) of the ten genes that remained significant after correcting for multiple testing are shown in figure 2.6. All the genes showed a significant upregulation in the SU ≥360 µmol/L + MSU crystals deposits group compared to the SU <360µmol/L group. While only *BIRC2, CXCL1, PANX1, TNFSF14, TXNIP* and *XIAP* showed a significant upregulation between the SU ≥360 µmol/L + MSU crystals deposits group compared to the SU ≥360µmol/L group. None of the genes showed a significant difference between the SU ≥360µmol/L and SU <360µmol/L groups.

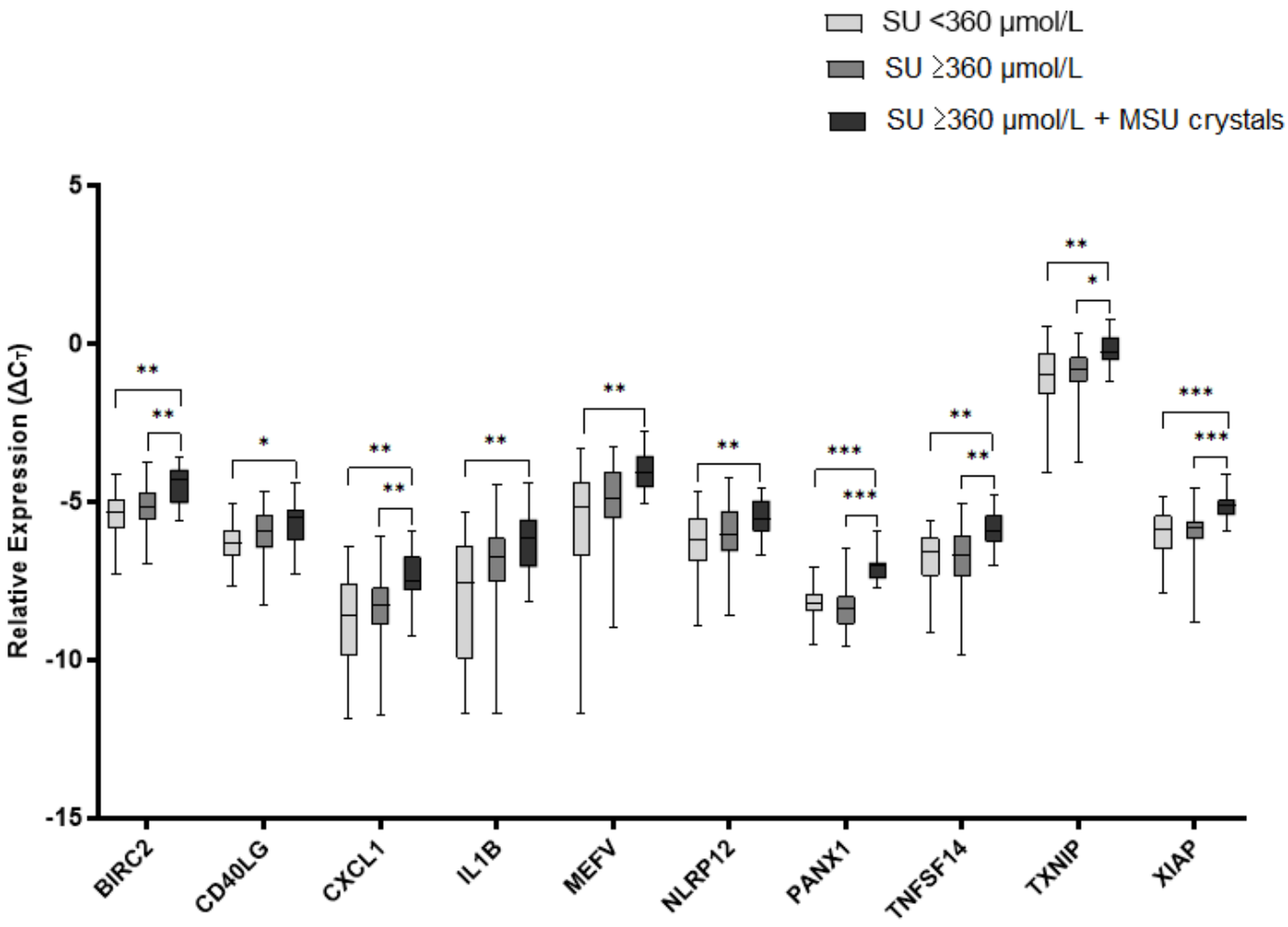


Figure 2.6 SOG: Box-plot of relative expression comparison of the eleven genes that showed a significant difference among all groups.

The Kruskal-Wallis H test was used to determine significance with Bonferroni post-hoc for pairwise comparisons. *p_{adj} <0.05; ** p_{adj} <0.005; ***p_{adj} <0.0001. SU, serum urate; MSU, monosodium urate

2.3.2 Gene expression profiling in patients with acute and intercritical gout

2.3.2.1 REACT study: Sample processing and demographic characteristics

A total of 21 participants met the entry criteria, and 19 gave consent for blood sample. However, only 16 provided blood during at least the baseline and one of the follow-up visits. In total, 51 blood samples were collected, from which total RNA was extracted. Quality control was analysed and 47 RNA samples met the parameters to undergo the cDNA synthesis.

Demographic characteristics and clinical information of the 16 participants whose samples were available for at least two time points, are summarised in Table 2.3. The mean (SD) age, BMI, and SU were 65.1 (14.1) years, 30.6 (4.7) kg/m^2**,** and 392.9 (113.6) µmol/L respectively. Gout's specific information was recorded, with the mean (SD) age of onset of 54.4 (16.8) years, and 3.7 (5.2) flares during the previous 12 months. The prevalence of people on allopurinol and febuxostat was 18.8% and 6.3% respectively.

Table 2.3 Demographic characteristics and clinical data of the REACT study participants

	Total
	n=16
Gender, male, n (%)	13 (81.25)
Age, years, mean (SD)	65.1 (14.1)
BMI, kg/m^2, mean (SD)	30.6 (4.6)
SU, µmol/L, mean (SD)	392.9 (113.6)
No. Flares, mean (SD)	3.7 (5.2)
Gout's age of onset, years, mean (SD)	54.4 (16.8)
Proportion on ULT, n (%)	
Allopurinol	3 (18.8)
Febuxostat	1 (6.3)
Comorbidities, n (%)	
Hypertension	9 (56.3)
Hyperlipidaemia	5 (31.3)
Diabetes	4 (25.0)
Heart attack	5 (31.3)

BMI: Body mass index; SU: Serum urate; ULT: Urate lowering treatment.

2.3.2.2 REACT study: Gene expression profiling

Although participants from the REACT study attended three visits: one during the gout flare, one at week 6, and one at week 12 after the baseline visit, only 11 individuals provided blood for RNA extraction during all the study visits. Moreover, a preliminary statistical analysis using Friedman test to compare repeated measures showed that there was no significant difference in any of the 86 genes between week 6 and week 12. Therefore, in order to increase the sample size, the intercritical gout stage was merged into one group by using the amplification data from week 12 if data from week 6 were missing. Thus, relative expression among groups was determined using the Wilcoxon signed-rank test, with a 5% FDR to correct for multiple testing. A fluorescence threshold of 0.3 was set for all the

expression arrays, and the amplification data was normalised to *RPLP0* housekeeping gene.

Out of the 86 genes under study, six (*CFLAR, NAIP, NFKBIA, NLRC4, NLRP6,* and *TLR2*) had a fold-change < 0.66 (corresponding to a fold-regulation ≤ -1.5) for the intercritical gout group compared to the active flare group, and a p value <0.05 (Figure 2.7 and Table 2.4). The greatest fold-regulation was observed for *NLRP6* with -2.38 (FC=0.42), followed by NAIP and TLR2 with -2.22 (FC=0.45) and -1.96 (FC=0.51) respectively. However, after correcting for multiple testing, none of the differences remained significant.

Table 2.4 Fold changes in gene expression for Intercritical gout compared to Gout flares group.

Gene	Fold-change	p value
CFLAR	0.54	0.049
NAIP	0.45	0.010
NFKBIA	0.55	0.034
NLRC4	0.60	0.049
NLRP6	0.42	0.011
TLR2	0.51	0.036

Relative expression was compared using the Wilcoxon signed-rank test; and correction for multiple testing was conducted using a 5% FDR.

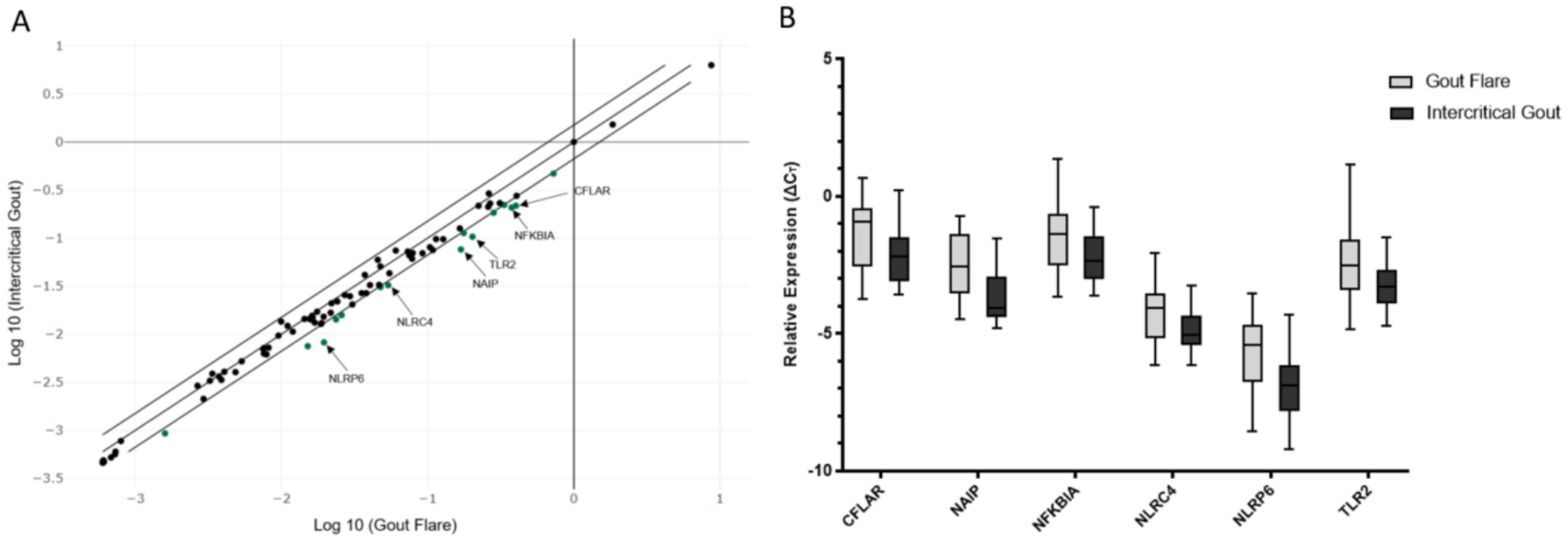


Figure 2.7 Gene expression of Gout Flare vs. Intercritical Gout graphs.

A) Scatter plot of the log base 10 ($2^{-\Delta CT}$) of each gene of the Human RT^2 Profiler PCR Array. The middle line indicates a fold change of 1.0, and upper and bottom lines indicate up and down regulation respectively. Green dots represent downregulated genes with a fold regulation ≤ -1.5. B) Box-plot of relative expression comparison of the six genes that showed a significant difference among groups. The Wilcoxon signed-rank test was used to determine significance; however, after 5% FDR correction, none of the differences remained significant.

2.3.3 Gene expression: comparison between patients with normouricaemia, asymptomatic hyperuricaemia and gout.

To have a complete scenario of the expression of the genes that showed a significant difference in the previous analyses, the relative expression of the 16 genes was compared between participants with gout flares and without symptomatic gout from REACT and SOG studies, respectively. The intercritical gout group from REACT study was not included in this analysis because these data corresponded to a repeated measure of the gout flare group. Although the SOG-AS study recruited participants with intercritical gout and blood samples were collected for gene expression profiling, the analysis of these samples was not completed due to Covid-19 pandemic restrictions.

Therefore, the expression of BIRC2, CD40LG, CFLAR, CXCL1, IL1β, MEFV, NAIP, NLRC4, NLRP6, NLRP12, PANX1, TLR2, TNFSF14, TXNIP and XIAP genes was compared among the following groups: SU <360µmol/L, SU ≥360µmol/L, SU ≥360µmol/L + asymptomatic MSU crystals and gout flare. Relative expression was compared using the Kruskal-Wallis H test with Bonferroni post-hoc test. Pairwise comparisons of the genes are shown in Figure 2.8, and fold changes are shown in Table 2.5 (Note this table includes the same fold changes for SU ≥360µmol/L and SU ≥360µmol/L + asymptomatic MSU crystals, compared to SU <360µmol/L, included in table 2.2). Out of the 16 genes, five showed a significant difference between the asymptomatic hyperuricaemia with MSU crystals group compared to the acute gout group. After correcting for multiple testing, CD40LG, PANX1 and TNFSF14 had lower mRNA levels in the acute gout group (mean rank = 30.06, 54.63 and 49.25, respectively) compared to the SU ≥360µmol/L + MSU crystals (mean rank = 72.78, 92.21 and 81.03, respectively). NLRC4 and NLRP12 had

higher mRNA levels in the acute gout group, compared to the normouricaemia or the SU ≥360µmol/L, but did not have a significant difference compared to the SU ≥360µmol/L + MSU crystals.

Table 2.5 Fold changes in gene expression for SU ≥360 µmol/L, SU ≥360 µmol/L + MSU crystals and gout flare, compared to SU <360 µmol/L.

	SU ≥360 µmol/L	SU ≥360 µmol/L + MSU crystals	Gout Flare	
Gene	**Fold-change**	**Fold-change**	**Fold-change**	**Adj. p value***
BIRC2	1.16	1.86	1.49	0.093
CD40LG	1.29	1.55	0.76	0.002
CFLAR	1.19	1.03	1.30	0.854
CXCL1	1.34	2.75	1.34	0.038
IL-1β	2.10	3.58	2.06	0.093
MEFV	1.67	3.10	1.40	0.033
NAIP	1.11	0.93	2.10	0.086
NFKBIA	1.05	0.96	1.82	0.183
NLRC4	0.82	1.46	2.31	0.0009
NLRP6	1.03	1.05	1.39	1.00
NLRP12	1.27	1.82	1.76	0.044
PANX1	0.91	2.20	1.09	0.0004
TLR2	0.96	1.07	1.92	0.167
TNFSF14	1.00	1.90	0.88	0.039
TXNIP	1.15	1.79	1.55	0.061
XIAP	1.00	1.79	1.25	0.0003

Relative expression was compared using the Kruskal-Wallis H test; and correction for multiple testing was conducted using a 5% FDR. *5% FDR adjusted p value. **NOTE:** This table includes the same fold changes for BIRC2, CD40LG, CXCL1, IL-1β, MEFV, NLRP12, PANX1, TNFSF14, TXNIP and XIAP shown in Table 2.2.

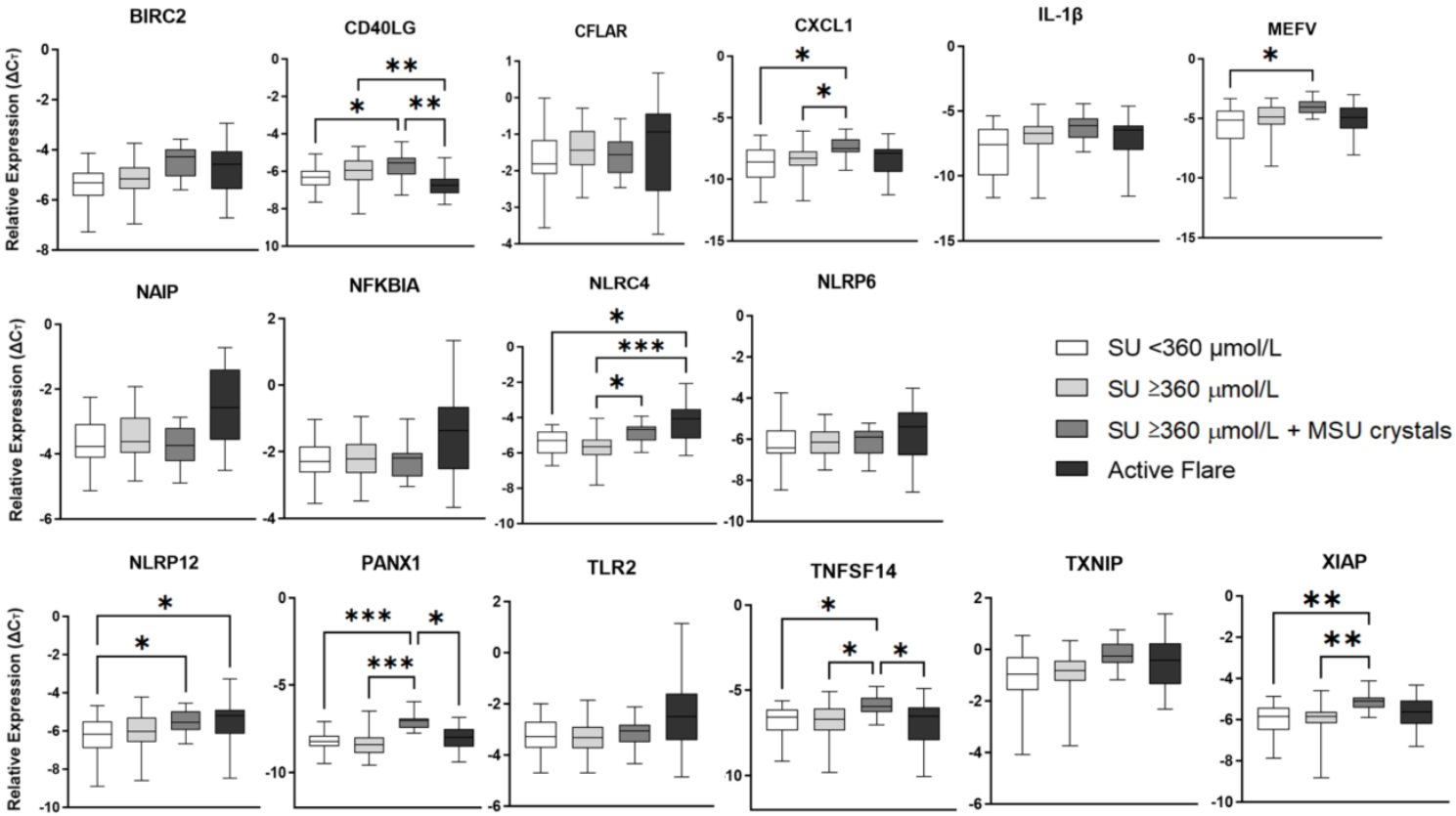


Figure 2.8 Box-plot of relative expression comparison of the sixteen genes that showed a significant difference amongst either SOG groups or REACT groups.

The Kruskal-Wallis H test was used to determine significance with Bonferroni post-hoc for pairwise comparisons. *p_{adj} <0.05; ** p_{adj} <0.001; ***p_{adj} <0.0001. SU, serum urate; MSU, monosodium urate.

2.3.4 Cytokine profiles in individuals with normouricaemia, hyperuricaemia without MSU crystal deposition, MSU crystal deposition without gout, intercritical gout and during a gout flare

2.3.4.1 Demographic characteristics

In total, 185 participants from SOG, REACT and SOG-AS studies were included in this analysis. Demographic data, comorbidities and gout-related information (where applicable) are shown in table 2.5. Participants from SOG study were classified into SU <360 µmol/L (n=33), SU ≥ 360 µmol/L (n=60), and SU ≥360 µmol/L + MSU crystals deposits (n=20). For SOG-AS and REACT studies, serum samples from the baseline visits were considered for this analysis; these corresponded to intercritical gout (n=54) and gout active flares (n=18), respectively. The mean (SD) age, BMI, and SU were 59.4 (13.8) years, 31.1 (5.0) kg/m^2**,** and 439.9 (79.8) µmol/L, respectively.

Table 2.6 Demographic characteristics and comorbidities data for participants from SOG, REACT and SOG-AS studies included in the cytokines measurement.

	SOG study				SOG-AS study	REACT study
	Total	**SU<360 µmol/L**	**SU≥360 µmol/L**	**SU≥ 360µmol/L + MSU crystal deposits**	**Intercritical gout**	**Gout flare**
	n=185	n=33	n=60	n=20	n=54	n=18
Gender, male, n (%)	181 (97.8)	33 (100)	60 (100)	20 (100)	53 (98.1)	15 (83.3)
Age, years, mean (SD)*	59.4 ± 13.8	48.3 ± 11.9	44.0 ± 10.5	46.8 ± 9.5	56.4 ± 12.1	69.6 ± 14.8
BMI, kg/m^2, mean (SD)*	31.1 ± 5.0	26.3 ± 6.1	27.7 ± 4.4	26.8 ± 3.2	31.0 ± 5.1	31.3 ± 4.7
SU, µmol/L, mean (SD)*	439.8 ± 79.8	309.2 ± 39.4	425.5 ± 42.9	423.3 ± 47.5	451.7 ± 62.1	397.9 ± 117.0
Comorbidities, n (%)						
Hypertension*	32 (17.3)	6 (18.2)	4 (6.7)	1 (5.0)	11 (20.4)	10 (55.5)
Hyperlipidaemia†	23 (12.4)	3 (9.1)	3 (5.0)	1 (5.0)	10 (18.5)	6 (33.3)
Diabetes†	9 (4.9)	1 (3.0)	1 (1.7)	0	3 (5.5)	4 (22.2)
Heart attack*	9 (4.9)	0	0	0	3 (5.5)	6 (33.3)
No. Flares, mean (SD)	3.3 ± 3.2	N/A	N/A	N/A	3.2 ± 2,3	3.7 ± 5.4
Gout's age of onset, years, mean (SD)	47.0 ± 14.0	N/A	N/A	N/A	45.1 ± 13.1	54.0 ± 15.5
Proportion on ULT, n (%)						
Allopurinol	21 (11.4)	N/A	N/A	N/A	18 (33.3)	3 (16.6)
Febuxostat	1 (0.5)	N/A	N/A	N/A	0	1 (5.5)

BMI: Body mass index; SU: Serum urate; MSU: Monosodium urate. N/A: Not applicable
P values were estimated by One-way ANOVA or independent t-test for continuous data, and chi-square test for categorical data.
*p value <0.001; † p value <0.01.

2.3.4.2 Cytokine profiles in individuals with normouricaemia, hyperuricaemia without MSU crystal deposition and MSU crystal deposition without gout, intercritical gout, and gout flares.

Serum samples, in aliquots of 2.0 mL, were sent to Affinity Biomarker Labs to measure the following cytokines: GRO-α, IL-1β, IL-6, IL-8, IL-18, IP-10, MCP-1, TNF-α, and VEGF-α measured using an MSD U-plex array, and hsCRP measured by latex agglutination test. Returned data were revised to replace values marked as out of the detection limit; these results were used to determine whether the levels of the cytokines of interest varied among groups.

First, cytokine levels were compared amongst the groups of REACT study to explore how the cytokines varied six and twelve weeks after the gout flare. This analysis was conducted using the Friedman test with Bonferroni post-hoc for pairwise comparison. Additionally, p values were corrected to account for multiple testing using a 5% FDR. Out of the 10 proteins under study, only VEGF-α and hsCRP showed a significant difference among groups (Figure 2.9). For VEGF-α, the rank difference among the intercritical gout-week 12 group (mean rank=48.0) was significantly lower than the gout flare group (mean rank=18.0; adjusted p value <0.0001), but the difference was not significant between week 6 group and gout flare group. For hsCRP, the difference for both intercritical gout week 6 (mean rank=22.0) and week 12 (mean rank=23.0) compared to the gout flare group (mean rank=45) were statistically significant (adjusted p values <0.0001 and p<0.001 respectively). There was no difference in any cytokines between week 6 and 12.

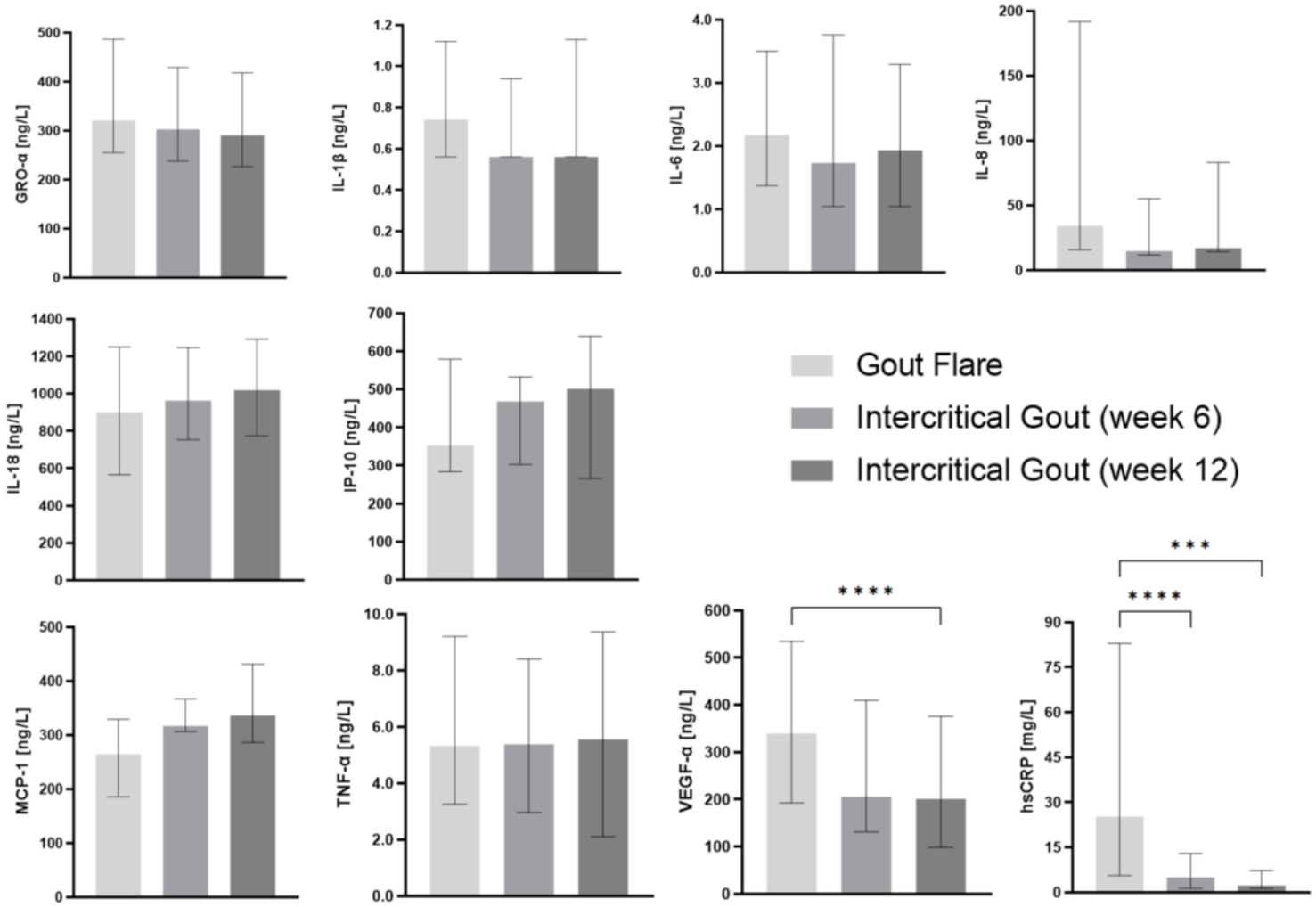


Figure 2.9 Protein levels of GRO-α, IL-1β, IL-6, IL-8, IL-18, IP-10, MCP-1, TNF-α, VEGF-α, and hsCRP in participants with gout flares and intercritical gout (6 weeks and 12 weeks after the flare).

Values represent median ± IQR. Differences among groups were compared using the Friedman test with Bonferroni post-hoc. *** p_{adj}<0.001; ****p_{adj}<0.0001.

Additionally, cytokines levels were also compared among the three groups of SOG study, the intercritical gout group of SOG-AS study, and the gout flare group of REACT study. The comparison was conducted using the Kruskal-Wallis H test with Bonferroni post-hoc, and p values were corrected for multiple testing using a 5% FDR. Results for all cytokines are presented in figure 2.10. All cytokines showed an adjusted p value <0.05, but when looking at pairwise comparisons, the differences were only significant between specific group pairs. For instance, only IL-1β, MCP-1 and hsCRP showed a significant difference among the intercritical gout group and the gout flare group. However, for IL-1β and MCP-1 the intercritical gout group had a mean rank (132.1 and 124.5, respectively) greater than the gout flare group (89.44 and 70.94, respectively), while for hsCRP the mean rank of the intercritical gout group was lower than the gout flare group (111.8 vs. 166.3). Interestingly, out of the 10 cytokines, 8 (GRO-α, IL-1β, IL-6, IL-18, IP-10, MCP-1, TNF-α, and hsCRP) showed significant differences among the intercritical gout group and at least one of the SOG study groups.

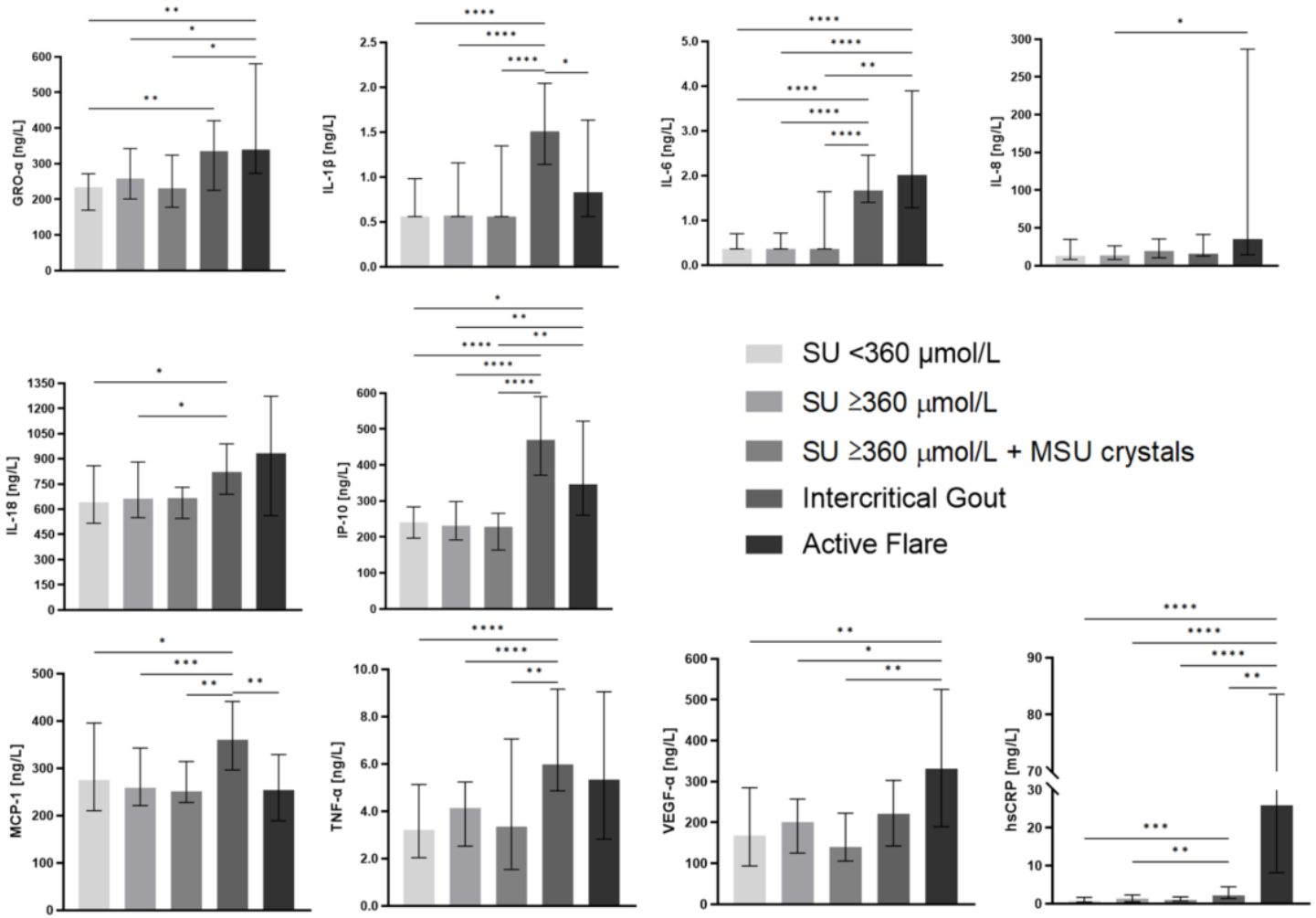


Figure 2.10 Protein levels of GRO-α, IL-1β, IL-6, IL-8, IL-18, IP-10, MCP-1, TNF-α, VEGF-α, and hsCRP in participants with SU <360 µmol/L, SU ≥360 µmol/L, SU ≥360 µmol/L with MSU crystal deposits, intercritical gout, and gout flares.

Values represent median ± IQR. Differences among groups were compared using the Kruskal-Wallis H-test with Bonferroni post-hoc. * p_{adj}<0.05; ** p_{adj}<0.01; *** p_{adj}<0.001; ****p_{adj}<0.0001.

2.4 Discussion

Two of the main challenges in knowledge about gout continues to be the understanding of the mechanisms involved in the transition from asymptomatic hyperuricaemia to gout, and the inflammatory profile of individuals with intercritical gout as a mean to understand the heterogeneity of gout symptoms. Given that it is well known that the NLRP3 inflammasome and TLRs 2-4 mediate the immune response to MSU crystals (Kingsbury et al., 2011; Major, Dalbeth, Stahl, & Merriman, 2018; Y. Shi et al., 2009), the objective of this chapter was to analyse the expression of genes and the levels of cytokines involved in those pathways in different sample sets, as outlined in figure 2.1.

Below is a discussion per section of the main findings and implications of each comparison.

2.4.1 Differential gene expression in participants with normouricaemia, asymptomatic hyperuricaemia and gout

In the analysis of asymptomatic hyperuricaemia vs. normouricaemia, out of the 86 genes included in the array, 10 (*BIRC2, CD40LG, CXCL1, IL-1β, MEFV, NLRP12, PANX1, TNFSF14, TXNIP*, and *XIAP*) showed a significant upregulation in the group of participants with hyperuricaemia and MSU crystal deposits, compared to the normouricaemia group. From these genes, only 6 (*BIRC2, CXCL1, PANX1, TNFSF14, TXNIP*, and *XIAP*) showed a significant difference among the hyperuricaemia with MSU crystal deposits and the hyperuricaemia only group. As described in Chapter 1, MSU crystals in joints are recognised as damage associated molecular patterns (DAMPs) by TLRs, which activate the MAP Kinase pathway and the translocation of the NF-κB transcription factor to initiate the expression of inactive pro-IL-1β, pro-IL18, and

NLRP3 component. The ingestion of MSU crystals also activate different signals that are required for the assembly of the NLRP3 inflammasome and the subsequent synthesis of caspase-1 that cleaves the inactive interleukins, to produce the active cytokines that are excreted from the macrophages (Jo et al., 2016; Kingsbury et al., 2011; Liu-Bryan et al., 2005). Several genes that were upregulated in this study play important roles in different mechanisms of the immune response to MSU crystals. These mechanisms can be divided into four: (1) TLRs and NF-κB downstream signalling, (2) NLRP3 inflammasome assembly mechanisms, (3) TLRs and NLRP3 inflammasome effector mechanisms, and (4) negative regulators. To facilitate the understanding, figures 2.11 and 2.12 give a visual representation of the inflammatory pathways initiated upon MSU deposition and the specific function of each gene that was found upregulated in this analysis. Figure 2.11 highlights in lilac the genes *BIRC2* and *XIAP* that are involved in downstream signalling of TLRs; in red the genes *PANX1* and *TXNIP* that participate in the signals required for the NLRP3 inflammasome assembly; and in blue the genes *CXCL1* and *IL-1β* that result as a consequence of the activation of TLRs and the NLRP3 inflammasome.

BIRC2 and *XIAP* are important components in the regulation of apoptosis by modulating both the NRLP3 inflammasome activity and the transcription factor NF-κB. *BIRC2* can provide positive and negative regulatory controls over NF-κB signalling via the canonical and non-canonical pathways respectively, which consequently alters the expression of pro-inflammatory cytokines (Tan et al., 2013). A study by Tseng *et al* reported that hypermethylation in *BIRC2* supressed its expression and exacerbated gout symptoms by increasing IL-1β production (Tseng et al., 2021). Although not directly linked to gout, several studies have analysed the effects of XIAP deficiency, and observed a sustained elevation of serum IL-1β and IL-18 (Lawlor et al., 2017; Yabal et al., 2014).

As previously described, the NLRP3 inflammasome assembly depends on three different signals. One of them is potassium efflux, which is controlled by pannexin-1 (*PANX1* gene) -a gap membrane channel- that together with P2RX7 mediate intracellular calcium and potassium concentrations in macrophages (Petrilli et al., 2007). It is known that probenecid, one of the available drug treatments for gout, inhibits anion transporters to avoid renal uptake of uric acid. Yet, it has been found that probenecid also inhibits pannexin-1. Although the exact mechanisms by which this happens are still unclear, as pannexin-1 has not been related to urate transporters, Silverman *et al* hypothesised that probenecid not only helps with urate lowering, but it can also cease inflammation by preventing the inflammasome activation (Silverman, Locovei, & Dahl, 2008). A second signal required for the inflammasome assembly is the presence of reactive oxygen species (ROS). An increase in ROS causes TRX to dissociate from TXNIP, which interacts with NLRP3 to allow its assembly with the remaining components of the inflammasome (Tschopp & Schroder, 2010). Dinesh & Rasool evaluated mRNA levels of several genes, included *IL-1β, TXNIP* and *NLRP3*, in an animal model of MSU induced inflammation to test the effect of berberine (an alkaloid used as food supplement). They observed that upon MSU crystals stimulation, *TXNIP* was upregulated, but when the macrophages were treated with berberine, the mRNA levels significantly decreased –similarly to what occurred when treated with colchicine-, suggesting TXNIP as not only as an essential component the immune response mediated by inflammasomes, but also as an important target for treatment approaches. (Dinesh & Rasool, 2017)

With regards to the effector mechanisms of TLRs and NLRP3 inflammasome activation, one of the hallmark signs of the inflammatory response to MSU deposits is the release of several cytokines, such as IL-1β, IL-18 and CXCL1.

In this study, the comparison of mRNA levels for *IL-1β* and *CXCL1* showed two of the greatest fold changes (3.79 and 2.90, respectively). This is interesting because the release of cytokines is the mechanism by which the immune response amplifies, and initiate the recruitment of macrophages and activation of leukocytes (A. Scanu et al., 2015; Scott, Ma, Viriyakosol, Terkeltaub, & Liu-Bryan, 2006; Y. Shi et al., 2009).

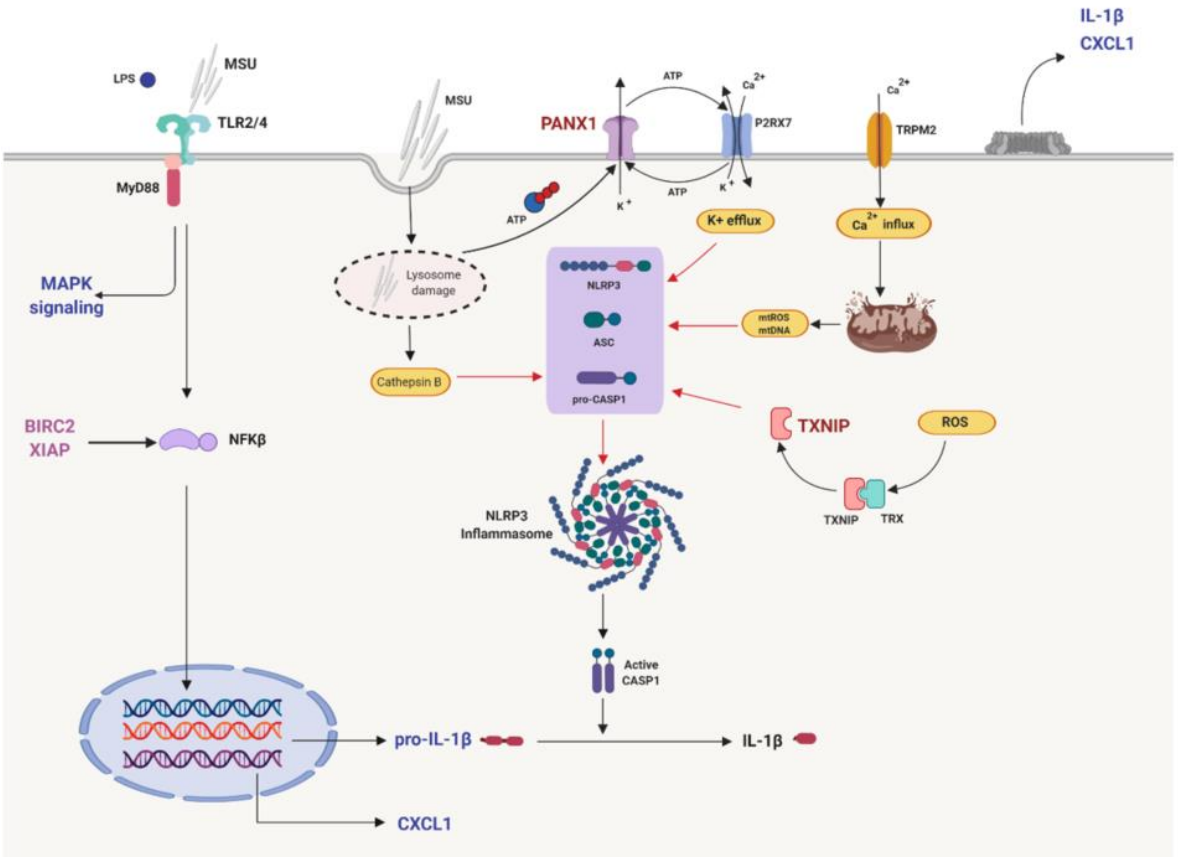


Figure 2.11. Molecular mechanisms of the inflammatory response induced by MSU crystals, highlighting eight genes that presented a significant upregulation in SOG analysis.

In lilac, *BIRC2* and *XIAP* that modulate the expression of the transcription factor NF-κB. In red, *PANX1* and *TXNIP* that contribute to the assembly of the NLRP3 inflammasome. In blue, *IL-1β* and *CXCL1* as the effector mechanisms of TLRs and NLRP3 signalling.

The genes mentioned above are related to pro-inflammatory mechanisms; however, this analysis revealed the upregulation of genes that are associated

to anti-inflammatory mechanisms as well. Figure 2.12 highlights in red the genes NLRP12 and MEFV that act as negative regulators in the TLRs signalling and the NLRP3 inflammasome assembly respectively.

First, the role of NLRP12 has been widely debated because it is classified as an inflammasome and is known to be involved in several bacterial infections, but the mechanisms that lead to its activation remain unknown (Vladimer et al., 2012). On the other hand, it has also been observed that NLRP12 works as a negative regulator of inflammation by inhibiting downstream signalling of TLRs and NF-κB, and consequently by limiting neutrophil recruitment to the site of inflammation (Tuladhar & Kanneganti, 2020). Loss of *NLRP12* has been associated to chronic colitis, hepatocellular carcinoma, and an increased expression of pro-inflammatory cytokines during viral and bacterial infections (Allen et al., 2013; L. Chen et al., 2017; Udden et al., 2019).

MEFV has also been identified as a negative regulator of inflammation, by preventing the assembly of inflammasomes that contain a pyrin domain (PYD), such as NLRP3. MEFV binds to the adaptor protein ASC, which is comprised by a PYD and a CARD domain, hindering the interaction with NLRP3 and pro-caspase-1 (Papin et al., 2007). Mutations in *MEFV* gene, responsible for the familial Mediterranean fever disease, alter the pyrin structure and block the interaction with ASC. This results in an increased activity of NLRP3 inflammasome and IL-1β production (Jamilloux et al., 2018). Balkarli *et al* investigated the association of these mutations with severity of symptoms in gout patients, and observed that heterozygous carriers of MEFV SNPs presented more flares per year, and the incidence of tophus was higher (Balkarli, Tepeli, Balkarli, Kaya, & Cobankara, 2018)

CD40LG and *TNFSF14* also showed a significant upregulation; however, their role in inflammasome-mediated immune response is less clear. Therefore, further investigation is needed to determine possible associations with MSU induced inflammation.

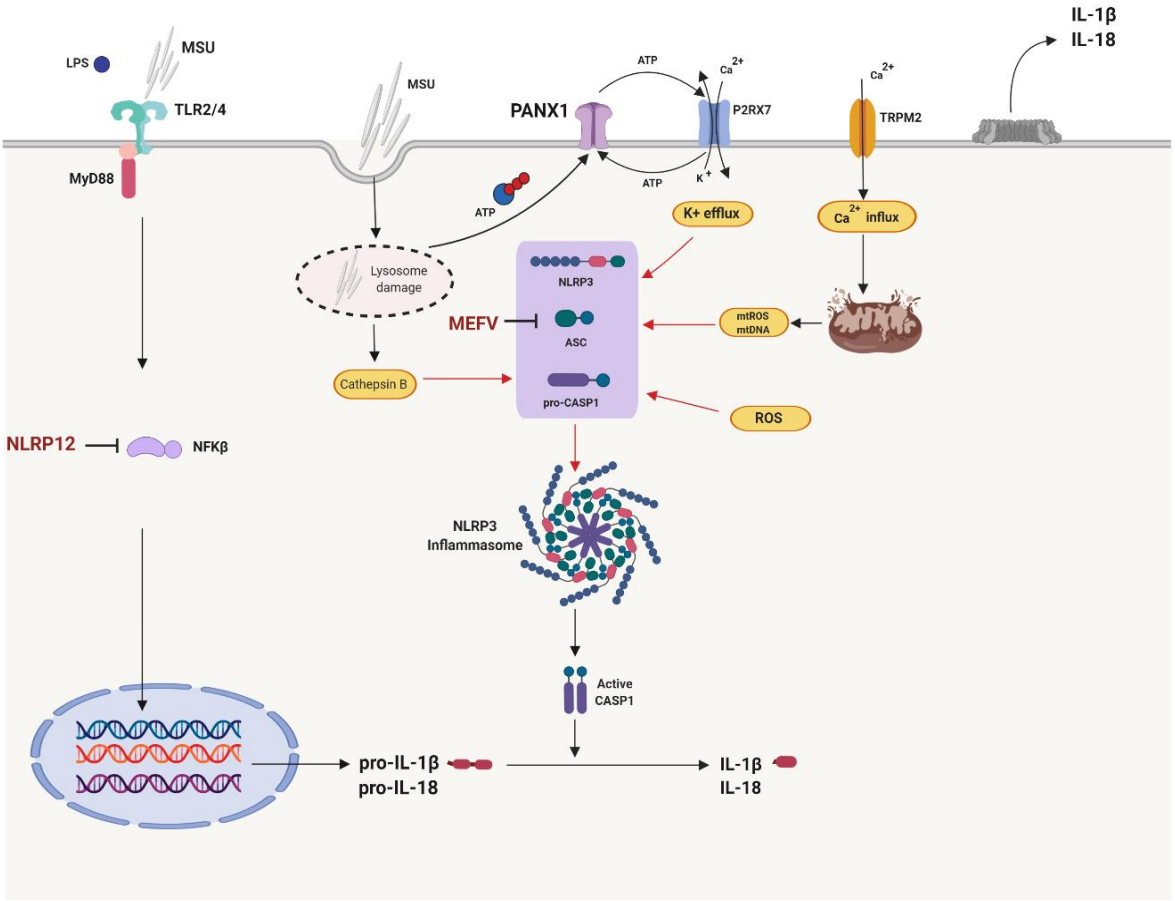


Figure 2.12. Molecular mechanisms of the inflammatory response induced by MSU crystals, highlighting two additional genes that presented a significant upregulation in SOG analysis.
In red, *NLRP12* and *MEFV* that inhibit the NF-κB cascade and the assembly of the NLRP3 inflammasome, respectively.

As discussed, some of these genes code for proteins that work as pro-inflammatory molecules, while others inhibit certain steps that could work as anti-inflammatory mechanisms. It is possible that this dynamic balance between

activation and inhibition of downstream signals upon recognition of MSU crystals prevents gout flares occurrence. This type of immune balance has been extensively described in infectious diseases like tuberculosis, H pylori infection, and hepatitis C (Cicchese et al., 2018). The pathogens causing the disease, when recognised by immune cells, activate the TLRs and NLRP3 inflammasome cascade, notwithstanding they can remain asymptomatic for extended periods (Larussa, Leone, Suraci, Imeneo, & Luzza, 2015; Negash, Olson, Griffin, & Gale, 2019; Wawrocki & Druszczynska, 2017). Moreover, although the immune response can be detected by measuring pro-inflammatory cytokines, it is not detrimental to the host's health. This hypothesis should be cautiously assessed and interpreted, because of the differences in the mechanisms that pathogens initiate, that DAMPs (such as MSU crystals) would not. However, supplementary research of immune balance in asymptomatic individuals in follow-up studies could contribute to the field.

The second analysis involved exploring gene expression in patients with gout, who attended a baseline visit during a gout flare and were followed-up 6 and 12 weeks later to compare the expression during the intercritical stage. Out of the 86 genes, six of them (*CFLAR, NAIP, NFKBIA, NLCR4, NLRP6* and *TLR2*) showed a significant downregulation during the intercritical period compared to the acute stage, but after correcting for multiple testing the differences did not remain significant. This could have happened due to the small sample size (n=16) of the study, or the fact that they were already treated with corticosteroids by the time the sample was collected.

A variety of studies have evaluated the clinical and biochemical features during the intercritical stage of gout patients. This has been done using imaging tools,

such as ultrasound and dual-energy computed tomography (DECT) to determine the persistence of MSU crystals (Breuer, Bogot, & Nesher, 2016; Das et al., 2017; Q. Wang et al., 2018), or by exploring cytokine profiles compared to those from healthy controls and from patients with gout flares (Diaz-Torne et al., 2021; Kienhorst et al., 2015). This is important not only because it allows the clinical and molecular differentiation of the characteristics of each stage, but also because it can lead to the identification of useful tools for the diagnosis of patients with suspected gout.

From the six genes that were upregulated in the intercritical gout group, two of them code for proteins that are central to the immune response induced by MSU crystals (Figure 2.13). To start with, TLR2 is one of the surface proteins that recognises MSU crystals and initiate the recruitment of MyD88 that activates the MAPK and NF-κB signalling pathways (Akira & Takeda, 2004). NFKBIA on the other hand, forms a dimer with NFKB1, that upon stimulation, is phosphorylated and dissociates to allow NF-κB translocation to the nucleus to start the transcription of pro-inflammatory cytokines (Mathes, O'Dea, Hoffmann, & Ghosh, 2008).

NAIP and *NLRC4* genes were also downregulated in this analysis. This is interesting because their encoded proteins assemble to form a different type of inflammasome. The NAIP/NLRC4 inflammasome has been described as one of the several innate immune responses that grants defence against bacterial infections (Kaiwen W. Chen et al., 2014; Franchi et al., 2012; Zhao et al., 2011), and information about its interaction with other inflammasomes under certain stimuli is emerging. For instance, Qu *et al* reported the interaction of NLRP3 with NLRC4 in *Salmonella typhimurium* infection that causes an amplified caspase-1 response. They suggested that the NAIP/NLRC4 interaction could

function as an additional signal that together with potassium efflux, cathepsin B synthesis and presence of ROS, allow the assembly of the NLRP3 inflammasome (Qu et al., 2016). Interestingly, a recent study by Teodoro-Braga *et al* found that murine NAIP expression was associated with an increased synthesis of IL-1β upon soluble serum uric acid stimulation (Braga et al., 2021).

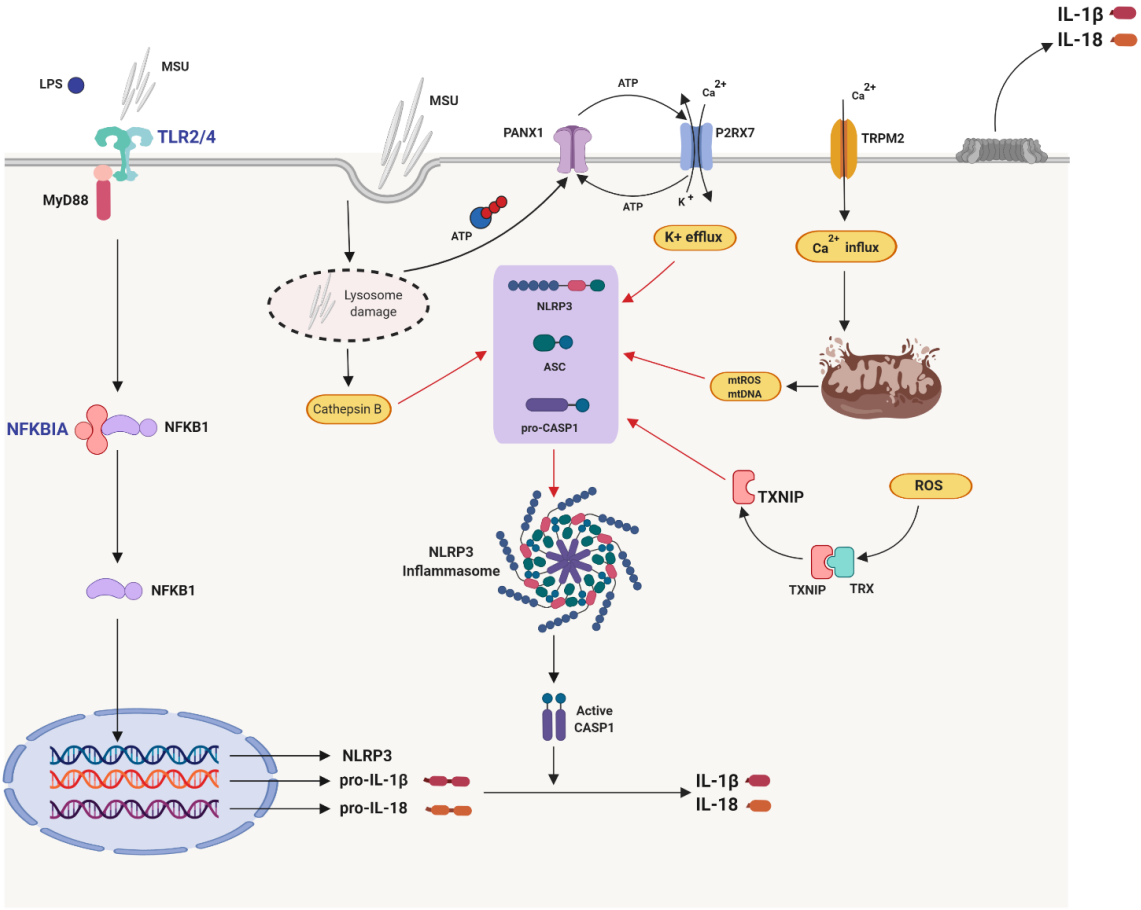


Figure 2.13. Molecular mechanisms of the inflammatory response induced by MSU crystals, highlighting two genes that presented a significant downregulation in REACT analysis.
In blue, *NFKBIA* and TLR2 that are crucial for the response to MSU crystals and the activation of the NF-κB signalling.

NLRP6 forms another inflammasome complex that has been associated to tissue specific functions. It plays an important role in gut inflammation and microbiota control, but its expression mechanisms are still to be elucidated

(Venuprasad & Theiss, 2021). CFLAR on the other side participates in the JNK/MAPK pathway, which is upstream of the NF-κB signalling, and is thought to negatively regulate inflammation and apoptosis. Several studies have described an important role of CFLAR in the improvement of cerebral injury and liver associated diseases (Y. Liu, Yu, & Chen, 2018; Xiaohong, Jun, Hongmei, & Fan, 2019). Whether NLRP6 and CLFAR activate during the response to MSU crystals or contribute to the resolution of gout flares, respectively, has not been reported and require additional research.

Although there are studies that have compared gene and protein levels in participants with gout during a flare or during the intercritical period, hitherto none of them has done this on the same participants as a follow-up study.

Finally, from the genes that showed significant differences between SOG and REACT groups separately, the last analysis consisted in exploring the expression of the sixteen genes in the whole spectrum of hyperuricaemia to gout flare. CD40LG, PANX1 and TNFSF14 had a significant downregulation during the acute flare compared to the group with asymptomatic MSU crystal deposition. CD40LG and TNFSF14 regulate T-cell proliferation and cytokine production, but the exact mechanisms have not been fully established (Giles et al., 2018; Mikolajczak et al., 2004). These findings were opposing to what we expected; however, the analysis was limited in terms of the sample size and the inclusion of an additional group of participants with intercritical gout. This work is currently underway and is expected to provide a clearer idea of whether the differences between these genes remain significant.

2.4.2 Cytokine measurements

As the mRNA levels for several genes were different among the groups under study, it was important to know the cytokine levels of those that were either found to be differentially expressed in these comparisons, or that play an important role in the immune response induced by MSU crystals. Therefore, 10 proteins (GRO-α, IL-1β, IL-18, IL-6, IL-8, IP-10, MCP-1, TNF-α, VEGF-α, and hsCRP) were measured and compared among the five groups.

First, in the comparison of the follow-up study REACT, among the gout flare and the intercritical stage (6 and 12 weeks after the flare), only VEGF-α and hsCRP showed a significant difference. At the time the mRNA analyses were conducted, VEGF-α had not been associated with gout. However, in 2019 the largest GWAS for SU and gout was published, and VEGF-α was reported as a risk loci (Adrienne Tin et al., 2019). Therefore, it was included in the protein measurements. VEGF-α is an angiogenic growth factor that has been related to a wide variety of diseases, including cancer (associated to tumour progression), diabetes, atherosclerosis, heart disease and retinopathies (Apte, Chen, & Ferrara, 2019). Angiogenesis has been described in other joint diseases such as rheumatoid arthritis, in which synovial capillaries increase to allow leukocytes migration during episodes of acute inflammation. Lioté *et al* analysed angiogenin (a potent angiogenic factor) levels in patients with different arthropathies, including gout, and observed that it was significantly elevated during the flare (LiotÉ, Champy, Moenner, Boval-Boizard, & Badet, 2003). It is possible that VEGF-α plays an important role in this process too, and as in rheumatoid arthritis, decreases with the resolution of the flare (Young Ho Lee & Bae, 2018).

hsCRP is the gold-standard biomarker of inflammation, as it increases up to 3000-fold during bacterial infections or tissue damage (Ansar & Ghosh, 2013). As expected, in this study there was a significant decrease during the intercritical period compared to the acute episode. Although useful for monitoring disease activity, hsCRP is not specific, thus not suitable on its own for diagnostic purposes. However, in chronic disorders such as obesity, diabetes, hypercholesterolemia and ischemic heart disease, CRP is usually persistently elevated and has been suggested as a risk predictor for these type of comorbidities (de Ferranti & Rifai, 2002; King, Mainous, Buchanan, & Pearson, 2003).

In order to determine if hsCRP and the pro-inflammatory cytokines included in this study, were also over expressed in acute and intercritical gout, compared to non-gout participants, an additional sample set of patients with intercritical gout (SOG-AS study) was included. This analysis showed that there was no variation among individuals with normouricaemia, hyperuricaemia and asymptomatic MSU crystal deposition; not even for IL-1β and GRO-α (CXCL1) that showed a differential mRNA expression. This is not surprising, since discordant effects on mRNA and protein levels have been debated for decades and may reflect counterbalance of pro and anti-inflammatory genes. These differences are likely to be a result of complex post-transcriptional and translational mechanisms (Edfors et al., 2016; Perl et al., 2017).

Conversely, out of the 10 proteins, the concentrations of GRO-α, IL-1β, IL-6, IL-18, IP-10, MCP-1, TNF-α, and hsCRP were significantly higher in the intercritical gout group compared to the non-gout participants from SOG study. This is interesting because it suggests that during intercritical gout there is evidence of a sub-clinical systemic inflammation. Several studies have reported similar

results after measuring pro-inflammatory cytokines (Diaz-Torne et al., 2021; Estevez-Garcia et al., 2018) or conducting ultrasound or DECT studies (Chowalloor, Raymond, Cheah, & Keen, 2020) that confirmed intra-articular inflammation even after the flare resolution. These observations are important due to the impact this may have on long term gout management by altering physician perception of gout, from an episodic illness with inactive crystal deposits in the intercritical phase to a chronic inflammatory illness. Adherence to treatment is known to be poor in gout patients, and is multi-factorial, but if sub-clinical systemic inflammation is persistent, long term ULT until crystal deposits are dissolved may be offered more emphatically by healthcare professionals.

2.4.3 General conclusions

In this research, significant differences in NLRP3 and TLRs associated pro-inflammatory genes, but not in cytokine levels, were found in asymptomatic hyperuricaemia individuals compared to those with SU <360µmol/L. Additionally, the results also indicated that during intercritical gout, there is evidence of systemic inflammation.

This research presents several limitations. First, participants with normouricaemia and hyperuricaemia were classified based on a single measurement of SU. Moreover, the sample sizes for several groups were limited. When we calculated the required sample size to detect a 2-fold change with an 80% power in pairwise comparisons, at least 26 samples per group were needed. However, the hyperuricaemia with MSU crystal deposits and gout participants (for the gene expression analysis) had 17 and 16 participants. With these numbers, we only reached a 50% power. Further validation in

independent and larger cohorts is recommended. Also, all the analyses were conducted in blood, therefore conclusions can only be driven in the context of systemic inflammatory mechanisms. Additional validation studies are required to determine if they replicate in synovial fluid resident lymphocyte and macrophage samples.

This study involved pooling samples from three different studies. Even though this allowed us to build a well defined larger set of samples that represent each stage from normouricaemia to gout, this presents several limitations. First, given the nature of each study, the recruitment criteria differed among studies. For instance, SOG study was limited to the recruitment of male participants, while REACT and SOG-AS recruited female and male participants. However, only 4 women took part of these studies. With these numbers (4 female vs. 181 male participants) we could not evaluate whether the differences in gene expression and cytokine measurements were sex-specific. Further studies with a heterogeneous sample are needed to examine if the mechanisms and genetic associations are sex-specific.

Also, the demographics and clinical characteristics (i.e. prevalence of comorbidities, previous treatment, etc.) were significantly different among groups, specifically among the non-gout vs. gout groups, and we did not adjust for these differences. Further analyses, such as MANCOVA, are suggested to determine whether differences in gene expression remain significant after adding covariates, such as age, gender, BMI, comorbidities, etc.

Additionally, although the protocols for collection and storage of blood samples were consistent for all studies, they were conducted in different time points, which may introduce variability. However, technical variability was minimised during subsequent methodologies (RNA extraction, cDNA synthesis, RT-PCR,

etc.) by limiting the processing of samples to one person, by using the same automatic lab equipment for all samples, and by introducing tight QC controls in each stage.

Finally, although valuable as a pilot study, the discovery power when using RT-PCR arrays is very limited, as it depends on the design and content of the plate. In this regard RNA sequencing can assist in finding novel transcripts and genetic variants contributing to gout.

Chapter 3. UK Biobank Genotyping and Polygenic Risk Score for gout.

3.1 Introduction

The introduction of personalised medicine has been explored for many years, and genetic studies have increased this possibility, with GWAS representing a patent impact to the field. While there is still an existing gap between GWAS loci and their functional consequences because the majority of risk variants are located in intronic regions, GWAS have contributed to the identification of pathways relevant to a disease, to the comparison of the genetic makeup among different populations and the assessment of risk prediction (Visscher et al., 2017). In gout, several GWAS have been conducted in different populations (Kottgen et al., 2013; C. Li et al., 2015; Nakayama et al., 2017). However, they have used smaller datasets, which has limited the discovery power and only few variants have reached genome-wide significance (e.g. the largest GWAS of gout included <2,500 gout cases (Kottgen et al., 2013)).

The most relevant insights in gout has been derived from GWAS looking at SU variation, given the causal role that hyperuricaemia plays in the development of gout. To date, over 180 variants have been associated with SU variation, and several studies have evaluated the correlation of these variants with gout. For instance, a study published by Phipps-Green et al investigated the association with gout of 28 SNPs that influence SU. They reported significant associations for eleven loci located in RREB1, TRIM46, SFMBT1, PRKAG2, A1CF, SFMBT1, VEGFA, PDZK1, IGF1R, MAF and HLF (Phipps-Green et al., 2016). Another study by Tin *et al* found significant associations for 55 SNPs, from which the urate transporter ABCG2, showed the largest effect (OR (CI) = 2.04 (1.96-2.12))

(Adrienne Tin et al., 2019). The results of GWAS of SU levels have been used also to generate predictive models of gout risk. Several studies have generated genetic risk scores using the sum of the effect sizes of genetic variants; however, they have used only the SNPs that reached genome-wide significance (Adrienne Tin et al., 2019; Yanfei Zhang & Lee, 2021). Polygenic risk scores (PRS) have acquired increasing interest in investigating the cumulative effects of genetic variants on a disease. Compared to other approaches, including those conducted to determine gout risk, PRS considers the inclusion of SNPs with small effects that even if not detected in GWAS, when added together would have a greater effect (Duncan et al., 2019).

Since one of the aims of genetic analyses would be its translation to a clinical setting, we wanted to construct a PRS model using gout GWAS summary statistics. For this purpose, we requested access to genotype data of the UK Biobank resource to generate a cohort of gout cases vs. controls. One of the crucial steps when conducting GWAS and PRS analyses is performing a robust QC of the genotyping data to exclude genotyping errors that would lead to type I and type II errors (Turner et al., 2011). Therefore, the aim of this chapter is to give a detailed explanation of the QC procedures, conducted by the UK Biobank and by ourselves to generate good quality working datasets for the PRS described in this chapter, and the analyses described in chapter 4.

3.2 Materials and Methods

3.2.1 UK Biobank cohort

This research was conducted using data from the UK Biobank resource (Project ID 45987. Appendix 1 includes the application for this PhD project). The UK Biobank is a large population-based prospective study that recruited ~500,000 participants, aged 40-69 years. Participants were recruited between 2006 and 2010 across England, Wales and Scotland. The baseline visit comprised electronic questionnaires and interviews, which aimed to collect data such as sociodemographic characteristics, lifestyle factors, health status, family history, cognitive function, etc. Physical and functional measurements were recorded, and biological samples (blood, urine and saliva) were also collected. Details about recruitment and data processing can be found in the key documents section on the UK Biobank website (www.ukbiobank.ac.uk).

Besides sociodemographic data, the outcomes and exposures of interest for this project included lifestyle information, such as smoking status and alcohol intake frequency, history of non-cancer illnesses and medications, genome wide genetic data, and biomarker measurements. Only participants who were classified as Caucasians (via self-report, followed by genetic grouping) and for whom genetic data were available, were included in the present study.

Information about the design of the genotyping arrays and their quality control (QC) methods, are described in detail elsewhere (Bycroft et al., 2017; Wain et al., 2015; Welsh, 2017). Sections 3.2.2-3.2.4 provide a brief summary of the UK Biobank internal QC.

3.2.2 The UK Biobank genotyping arrays

The data release included a total of 502,536 participants. However, genotyping data were only available for 488,377 participants. Genotyping was performed in two arrays. The UK BiLEVE Axiom array was used to genotype the first 49,950 samples that took part of the UK Biobank Lung Exome Variant Evaluation study (Wain et al., 2015), and included 807,411 markers. The remaining 438,427 samples were genotyped using the UK Biobank Axiom™ Array, which included 825,927 markers. The Axiom™ array content summary document is available in the UK Biobank website (https://www.ukbiobank.ac.uk/scientists-3/uk-biobank-axiom-array). Both arrays shared 95% of marker content. Figure 3.1 summarises the arrays content.

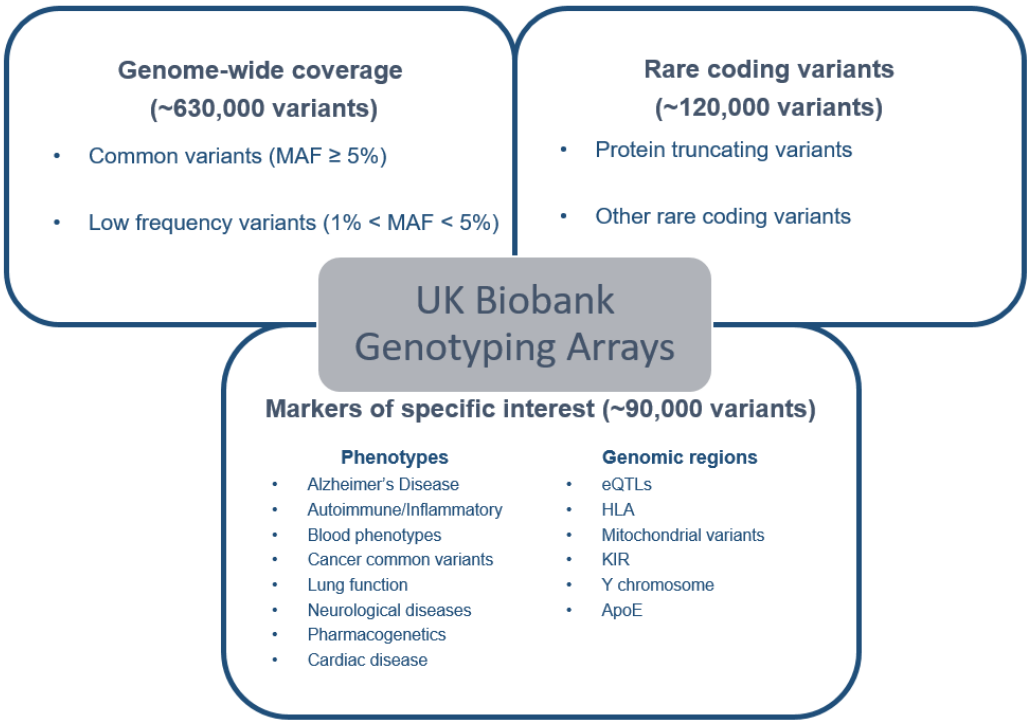


Figure 3.1 Summary of UK Biobank/Affymetrix arrays content.
This scheme describes the content by category, and an approximate number of markers within each category. (Bycroft et al., 2017)

3.2.3 DNA extraction and genotyping processing

Blood samples were collected during the baseline visit to the UK Biobank assessment centres. Samples for genotyping were selected, and DNA was extracted, quantified and analysed to guarantee acceptable quality control parameters. DNA samples were shipped to Affymetrix Research facilities in 96-well plates (each plate containing 94 DNA samples and 2 control samples from the 1000 Genomes Project) (Welsh, 2017; Welsh, Peakman, Sheard, & Almond, 2017).

Genotyping was conducted by Affymetrix in 106 batches of ~4,700 samples each: 11 batches were genotyped on the UK BiLEVE Axiom array, and 95 batches were genotyped on the UK Biobank Axiom™ Array. Fluorescence intensities of each allele were measured. This allows the formation of clusters (homozygous clusters –major or minor, and heterozygous cluster), from which genotypes can be inferred (example shown in figure 3.2).

Non-autosomal markers were analysed differently. The Y chromosome and mitochondrial markers were assessed by inspecting cluster plots, where only two genotype categories are expected. The pseudo-autosomal regions (PAR) of the X chromosome were assessed as autosomal markers.

Filters to exclude markers with poor clustering were applied per batch. A total of 35,014 markers did not meet the cluster criteria across batches, or showed complex cluster properties (multi-allelic markers); therefore, they were removed from the dataset. Detailed information regarding laboratory processing and genotype data generation by Affymetrix can be found in the online showcase of UK Biobank under the resource IDs 590 and 368 respectively (http://biobank.ndph.ox.ac.uk/showcase/).

Genotype calling was available for 812,428 markers in a dataset comprised by 489,212 individuals, on which further quality control procedures were applied.

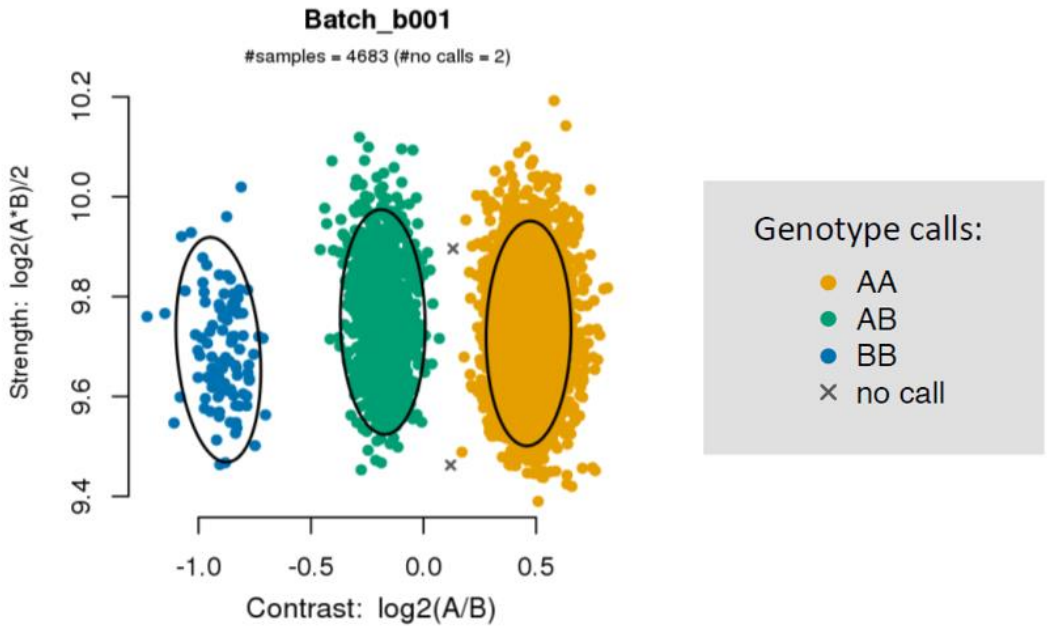


Figure 3.2 Example of clustering and genotype calling for one marker in Batch 001. X and Y axes represent transformations of the probes intensities, targeting allele "A" and allele "B". The ovals indicate the location and distribution for each cluster. Adapted from (Bycroft et al., 2017).

3.2.4 UK Biobank samples and markers Quality Control (QC)

Large-scale and strong population structure studies, such as the UK Biobank project, give rise to several issues that needed to be addressed prior to conducting any analysis. These characteristics, plus the fact that genotyping was performed using different arrays and many batches, required a robust QC pipeline to account for all the effects.

3.2.4.1 Marker-based QC

Given that conventional QC tests are ineffective when studies show ethnic diversity, marker-based QC was conducted exclusively using individuals with European ancestry (463,844 individuals). Selection of individuals with European ancestry was achieved by using principal component analysis (PCA) to account for population structure (described in section 3.2.4.2).

Four QC tests were designed and applied for each marker in each batch separately to explain: (1) Batch effects, (2) Plate effects, (3) Gender effects, and (4) Deviation from Hardy-Weinberg equilibrium (HWE). Because individuals were selected from the same ethnicity, differences in allele frequencies across batches, plates or gender were not expected. Therefore, significant differences were indicative of genotyping errors. To determine if such effects impacted allele frequencies, Fisher's exact test was used on the 2x2 contingency table for haploid markers, or on the 2x3 contingency table for autosomal (and PAR) markers. Deviation from HWE was evaluated on diploid markers using the exact test described by Wigginton et al (Wigginton, Cutler, & Abecasis, 2005).

A p-value of 10^{-12} was chosen as the cut-off value for all tests. Such a p-value was selected in order to remove only markers with strong deviation from the null hypotheses.

Additionally, two tests were applied for each marker across all batches. The first one was designed to correct for array effects. Fisher's exact test was used to determine whether individuals genotyped on the UK BiLEVE Axiom Array had the same allele frequencies as those genotyped on the UK Biobank Axiom Array. As for the previous four tests, if markers had a p-value less than the defined threshold (10^{-12}), they were removed from the dataset. Finally, as previously mentioned, all plates included DNA of two participants from the 1000

Genomes project as controls. A discordance metric was generated for each of the controls and each marker, and if a marker had a concordance less than 95% for at least one control, it was excluded from the dataset.

A total of 8,425 markers failed at least one of the QC tests, and were consequently removed from the released dataset.

3.2.4.2 Sample-based QC

In order to perform sample QC, only SNPs that passed marker quality control and were in both genotyping arrays were selected (621,641 autosomal and 15,766 non-autosomal).

Population structure: Genetic ancestry was captured by principal components analysis. PCA was conducted using the FastPCA algorithm described by Rokhlin et al (Rokhlin, Szlam, & Tygert, 2009). The analysis was done in two rounds, the first one aimed to identify unrelated individuals and to adjust for heterozygosity. The second-round involved computation of the top 40 PCs after samples were filtered according to their kinship coefficients, quality scores (call-rate and heterozygosity), and markers were pruned to minimise linkage disequilibrium (LD).

Gender mismatches: Discrepancies among self-reported gender and genetic gender can occur as a consequence of sample mishandling, chromosome aneuploidies, or gender dysphoria. Therefore, they should also be captured for QC purposes. Genetic gender was inferred from measured intensities of

markers on the X and Y chromosomes. For this PhD thesis, all individuals with putative aneuploidies were excluded.

Kinship: Like the previous parameter, the calculation of kinship coefficients is also essential for QC purposes, not only because it allows to identify sample mishandling (specifically experimental duplicates), but it also allows to identify genuine familiar relatedness that can affect allele frequencies. Identity by descent was calculated for all pairs of samples using the software KING, and kinship coefficients were recorded. The released dataset included a variable with individuals who were inferred to be third degree relatives or closer.

Call-rate and Heterozygosity: Raw heterozygosity was calculated using 605,857 autosomal SNPs and adjusting for population structure using the first round of PCA computation. Adjusted heterozygosity values for all participants were provided as part of the released dataset. Additionally, call-rate was calculated using PLINK. An individual was classified as an outlier if the adjusted heterozygosity was greater than 0.1903 and the call-rate was less than 95%.

Even though several samples did not meet the QC criteria, they were not removed from the released dataset, but variables to identify them were provided. Individual phenotype associated data were provided as Stata .raw files, while genotype data were provided as binary PED files. Therefore, StataSE 15 was used to manage phenotype data to filter individuals who did not meet the UK Biobank QC criteria, while PLINK was used to exclude the markers that failed the QC.

3.2.5 PLINK QC

Although the UK Biobank designed and applied QC tests for markers and samples, only those with strong deviations failed the QC criteria. Consequently, additional QC refinement was needed prior conducting further analyses. To reduce the computing time when performing the QC procedures, cases and controls for each analysis were extracted from the working dataset and additional QC filters were applied using PLINK version 1.9. PLINK is an open-source command line software, designed to efficiently manage and analyse whole-genome data (S. Purcell et al., 2007). The UK Biobank provided genotype data as binary PED files, which include *.bed files containing the genotype information, *.fam files containing individuals' data and *.bim files containing markers information (Figure 3.3).

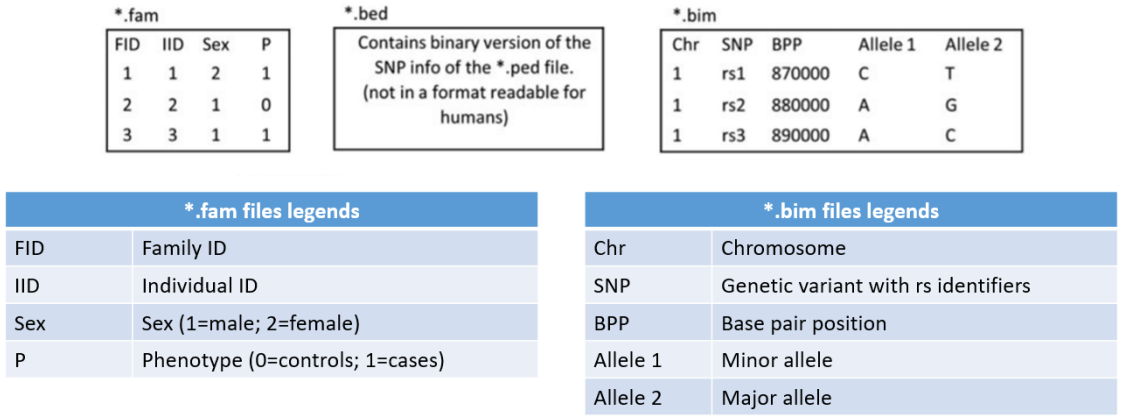


Figure 3.3 Overview of PLINK binary files.
Adapted from (Marees et al., 2018).

Filters to control for relatedness, population structure and gender mismatches were not necessary at this stage because the UK Biobank released data included variables to select individuals according to their ancestry, kinship coefficients and genetic sex. Therefore, the working dataset excluded individuals with non-European ancestry, third degree (or greater) relatives and sex mismatches. Table 3.1 explains each command of the PLINK QC filters, which included the following parameters:

Samples and markers call-rate: This parameter excluded individuals and SNPs with high levels of missingness, which can be indicative of poor DNA quality, technical errors and poor probe design. Individuals with a call rate <90% and markers with a call rate <95% were removed from the dataset.

Commands:

```
>plink --bfile FILE_1 --mind 0.1 --make-bed --out FILE_2
>plink --bfile FILE_2 --geno 0.05 --make-bed --out FILE_3
```

Hardy-Weinberg Equilibrium (HWE): Genotype and allele frequencies are expected to remain constant over generations, but this equilibrium can be disturbed by different factors, such as mutations, non-random mating and genetic drift. However, if genetic association studies are conducted in individuals from the same ethnicity, differences in allele frequencies are not expected, otherwise they can be a result of genotyping errors. Therefore, a more stringent evaluation was needed to identify markers deviating from HWE. PLINK generated a file that included the genotype counts and the HWE test statistics for common SNPs (minor allele frequency (MAF) >0.05). The SNPs

were subsequently filtered and those with a p value <$6.82x10^{-8}$ (Bonferroni corrected: 0.05/733138) in the control group, were removed from the dataset.

Commands:

```
>plink --bfile FILE_3 --maf 0.05 --hardy --out FILE.hwe
>plink --bfile FILE_3 --exclude hwe_SNPs.txt --make-bed --out FILE_4
```

Heterozygosity: The identification of heterozygosity outliers was the last step in the QC pipeline. This was important because high levels of heterozygosity can be an indicator of poor sample quality or cross contamination, while low levels of heterozygosity might be a result of inbreeding. Prior analysing heterozygosity coefficients, a pruning step was needed to extract common, independent and autosomal SNPs with the following parameters: window size = 50kb, step size = 5 SNPs, and correlation coefficient (r^2) = 0.2.

```
>plink --bfile FILE_4 --maf 0.1 --extract autosomal_SNPs.txt --indep-pairwise 50 5 0.2 --out FILE_indepSNP.prune.in
```

Inbreeding coefficients were then generated with PLINK. To identify the outliers, the mean and standard deviation (SD) were calculated, and those deviating ± 3SDs were excluded from the final dataset.

Command:

```
>plink --bfile FILE_4 --extract FILE_indepSNP.prune.in --het --out FILE.het
>plink --bfile FILE_4 --remove heterozygosity_outliers.txt --make-bed --out FILE_5
```

Table 3.1 PLINK commands for quality control procedure.

	Command	Description	Selected parameter
Data management	--bfile	Binary file name	FILE_N
	--make-bed	Creates a new PLINK binary file	-
	--out	Output name	FILE_N
	--exclude	Excludes SNPs from the input file	txt file with rs identifiers
	--extract	Extracts SNPs from the input file	txt file with rs identifiers
	--remove	Removes individuals from the input file	txt file with individual IDs
QC commands	--mind	Excludes individuals with low genotype calls	0.1
	--geno	Excludes SNPs with low genotype calls	0.05
	--maf	Filters SNPs according to a defined minor allele frequency	0.05 (for HWE), 0.1 (for pruning step)
	--hardy	Generates a file with genotype counts and Hardy-Weinberg statistics	-
	--indep-pairwise	Pruning step to identify independent SNPs	Window size = 50kb, step size = 5, r2=0.2
	--het	Computes inbreeding coefficients	-

3.2.6 Polygenic risk scoring

Quality controlled genotyping data was used to compute polygenic risk scores (PRS) for all individuals in the target dataset using PRSice version 2.0. PRSice is a command-line software package that allows the calculation, evaluation and plotting of PRS at a range of p-value thresholds to identify the best-fit model. It is written as an R code, compiled for bash data management in PLINK. PRSice requires GWAS results on a phenotype of interest that work as the base dataset, and genotype data as the input or target dataset. These datasets require the same extensive GWAS QC to control for markers and samples missingness, HWE and heterozygosity, which has been described in the previous section. (S. W. Choi & O'Reilly, 2019; Euesden, Lewis, & O'Reilly, 2014). Figure 3.4 summarises the PRS processing using PRSice. The following sections describe the datasets used for the generation of a PRS for gout vs. controls, and the parameters used to generate the model.

3.2.6.1 Base dataset

In 2013, Köttgen *et al* published the largest GWAS meta-analysis for gout. This included 14 studies, comprising 2,115 gout cases and 67,259 controls of European ancestry within the Global Urate Genetics Consortium (GUGC) (Kottgen et al., 2013). The summary statistics that resulted from this study were used as the base dataset for the calculation of PRS in people with gout. They are publicly available via the Genome-Wide Repository of Associations Between SNPs and Phenotypes (GRASP) database (Leslie, O'Donnell, & Johnson, 2014). The summary statistics were provided as *.csv files, which were converted and merged into a single *.txt file containing all the association analysis results for >2.5 million SNPs on gout.

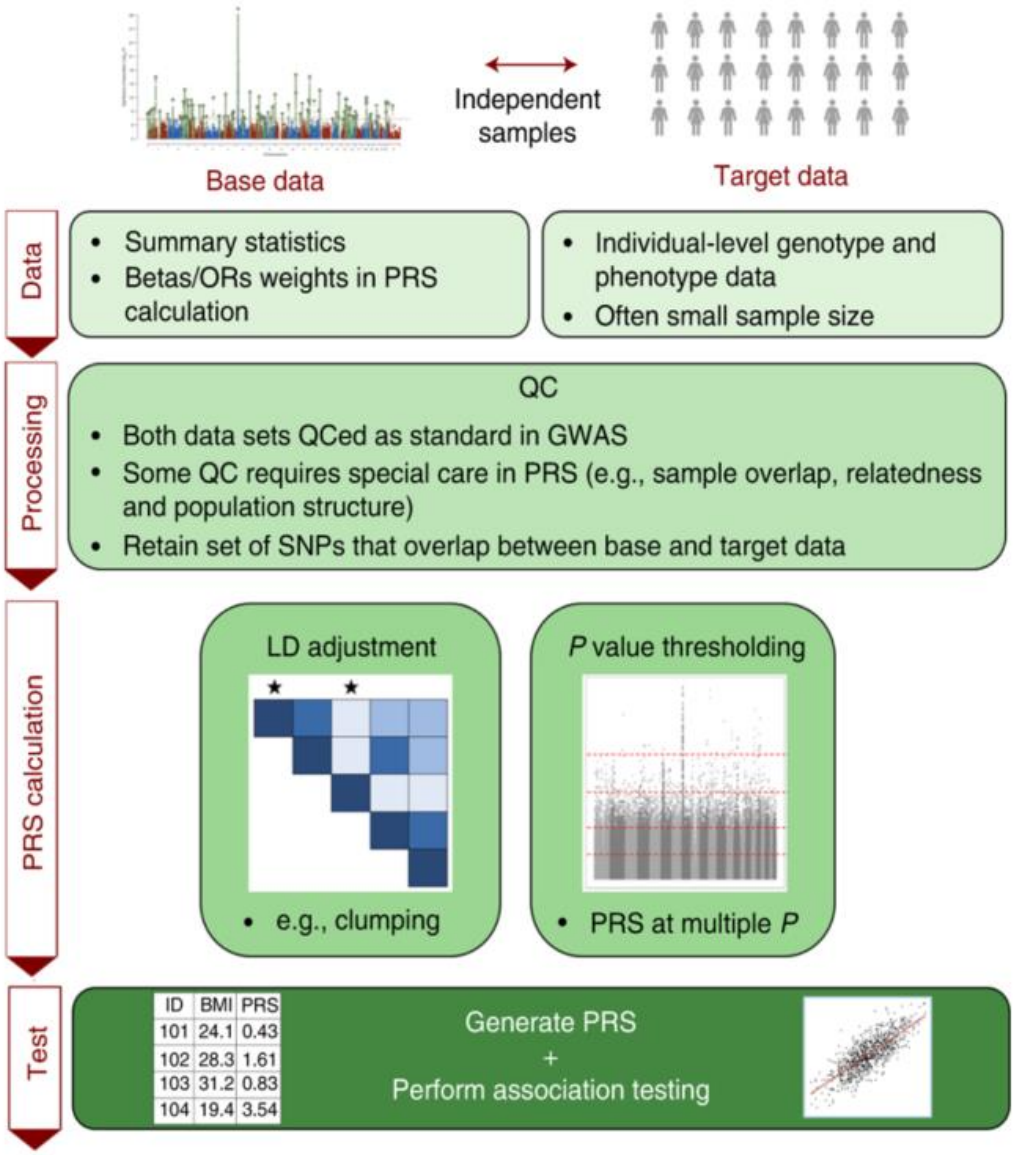


Figure 3.4 Polygenic Risk Score Pipeline.
Adapted from (S. W. Choi, Mak, & O'Reilly, 2020)

3.2.6.2 Target dataset: case-control definition

The target dataset was assembled from the UK Biobank genotyping data. Phenotype and genotype data of gout cases and controls were extracted using Stata and PLINK 1.9, respectively. Definition of cases and controls was as follows:

Gout cases: Gout cases were defined following three different criteria (Cadzow, Merriman, & Dalbeth, 2017): (1) Self reports, (2) use of urate lowering treatment (ULT) and (3) hospital diagnosis. For self-reports of gout, the data field "non-cancer illnesses" was used to extract participants who reported physician diagnosed gout at the baseline visit. Use of ULT consisted of self-reports of being on allopurinol, probenecid or sulfinpyrazone. Participants on ULT without gout diagnosis via self-reports or hospital records, and with either a primary or a secondary diagnosis of leukaemia (ICD-10 codes C90-C96) or lymphoma (ICD-10 codes C81-C88) were excluded. Finally, hospital diagnosis of gout was defined if participants had a hospital record of primary or secondary episodes, which were identified with the following ICD-10 codes: M10 (gout), M100 (idiopathic gout), M101 (lead-induced gout), M102 (drug-induced gout), M103 (gout due to impaired renal function), M104 (other secondary gout) and M109 (gout, unspecified). Figure 3.5 shows the gout cases identified per definition.

Controls: Individuals without gout from the UK Biobank were selected as controls to assemble the target dataset. Gender is recognised as a strong confounder in gout (Kuo et al., 2015), and because the UK Biobank is comprised by 53.4% of female participants, a random selection of controls would have led to an unequal distribution among gout cases and controls. Therefore, cases were matched to up to five controls based on gender.

These phenotype-genotype taget dataset were merged into a PLINK binary file, which underwent the QC procedures described in section 3.2.5.

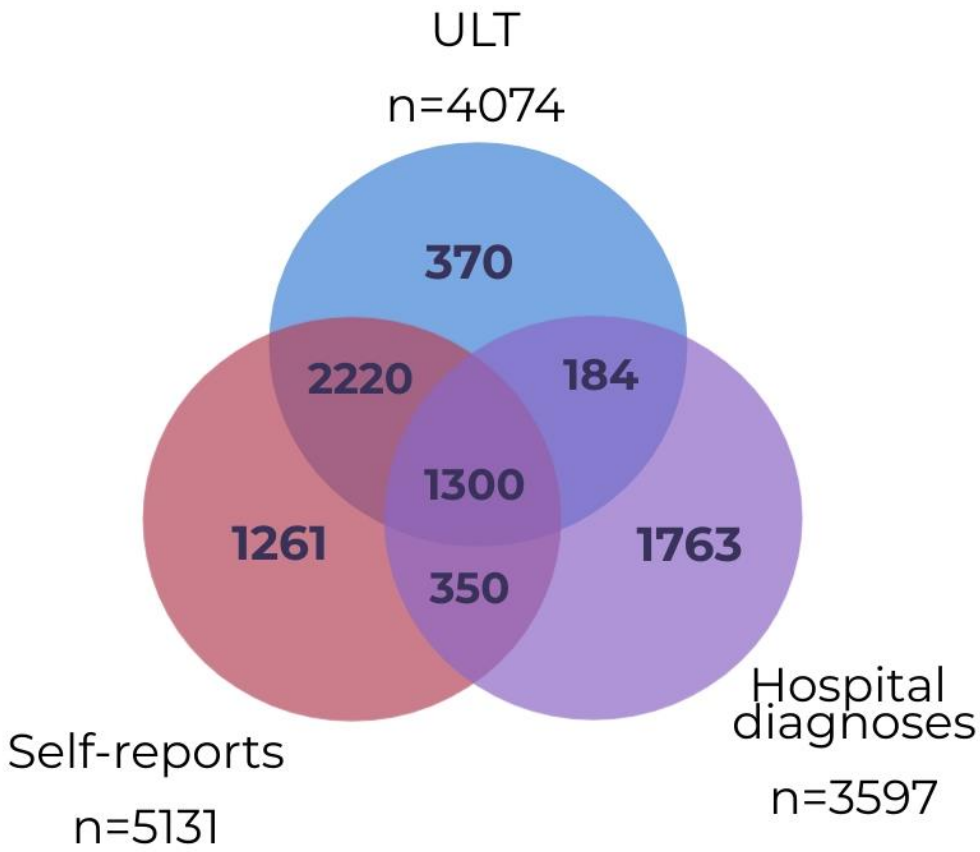


Figure 3.5 Venn diagram of gout cases from the UK Biobank, identified via self-reports, urate lowering treatment and hospital records.

3.2.6.3 PRSice script

PRSice allows setting of multiple parameters, such as LD clumping, selection of p-value thresholds, inclusion of covariates to compute individual PRS, use of an external panel to compute LD calculations, etc. Table 3.2 gives a description of all the commands and settings used to generate the PRS for this project Briefly, LD clumping was set to an r^2 >0.1 using a 500kb window; upper and lower p-value thresholds were set as 0 and 1, respectively, with increments of $1.0x10^{-6}$. The use of an external LD reference panel is suggested when sample sizes of the target dataset are small. Because this project involved >40,000 participants, the same target file was used as the LD reference panel.

Script:

```
Rscript PRSice.R --dir . \
--prsice ./PRSice \
--base GUGC_SummaryStats.txt \
--target Gout_matched_dataset \
--thread max \
--print-snp T \
--stat OR \
--binary-target T \
--no-clump F \
--clump-kb 500 \
--clump-r2 0.1 \
--clump-p 1 \
--perm 10000 \
--lower 0 \
--interval 0.000001 \
--upper 1 \
--out Gout_controls_PRS
```

Table 3.2 PRSice commands used to generate the PRS for gout vs. controls.

Command	Description	Selected parameter
--base	Base/Training dataset	Gout GWAS Summary statistics (Kottgen et al., 2013)
--target	Target dataset, binary PLINK file	Replication cohort
--thread	Number of thread	Maximum
--print-snp	Generates an output with the SNPs that remain after clumping	TRUE
--stat	Column header containing the effect size estimate	Odds Ratios
--binary-target	Indicates if the target dataset includes a dichotomous phenotype	TRUE
--no-clump	If true, PRSice do not compute LD clumping	FALSE
--clump-kb	Indicates the window for clumping	500 kb
--clump-r2	Indicates the r2 threshold for clumping	0.1
--clump-p	Indicates the p value threshold for clumping	1
--perm	Number of permutation to generate the empirical p value	10000
--lower	Indicates the starting p value threshold to be tested	0
--interval	Indicates the step size of the p value threshold	0.000001
--upper	Indicates the final p value threshold to be tested	1
--out	Output name	FILE_NAME

3.2.6.4 Statistical analysis

Odds ratios (OR) and *p* values from the GWAS summary statistics were used to calculate the best PRS model, which was generated from testing different *p* value thresholds. The best-fit model was defined by the largest Nagelkerke's R^2 value. Logistic regression was used to estimate the effect of the demographic variables for inclusion into the predictive models. To determine the predictive ability of the PRS, the demographics (which included age at recruitment, sex and body mass index (BMI)) and the combined model, the area under the receiving operatic characteristic curve (AUROC) was used. Logistic regressions and the AUROC were conducted using SPSS Statistics 24.

3.3 Results

3.3.1 UK Biobank Cohort: Demographic characteristics

The UK Biobank released dataset included 502,536 participants. After excluding participants without genetic data available, 488,288 individuals were selected for QC refinement, which resulted in 354,825 participants. Figure 3.6 summarises the QC filters for ethnicity, kinship, gender mismatches, call rate and heterozygosity. Demographic characteristics, lifestyle information and comorbidities data are summarised in table 3.3. The resulting UK Biobank cohort was comprised by 46.1% men. The mean (SD) age and BMI were 56.96 (7.91) years and 27.42 (4.76) kg/m^2, respectively. The prevalence for diabetes, hypertension, hypercholesterolemia, ischaemic heart disease and cardiac failure were 4.18, 26.64, 12.29, 4.51 and 0.07, respectively.

3.3.2 PRS Target dataset: Demographic characteristics

Gout cases and controls (matched by gender) were derived from the quality controlled UK Biobank cohort. These samples and markers underwent a second stage of QC filters in PLINK. 7,448 gout cases, 39,959 controls, and 717,091 genetic variants passed the QC parameters. These data integrated the target dataset for PRS calculations. It was comprised by 89.8% men, and their mean (SD) age, BMI and SU were 57.46 (7.94) years, 28.11 (4.53) kg/m^2 and 5.92 mg/dL, respectively. Table 3.4 summarises demographic, lifestyle and comorbidities data for cases and controls of the target dataset.

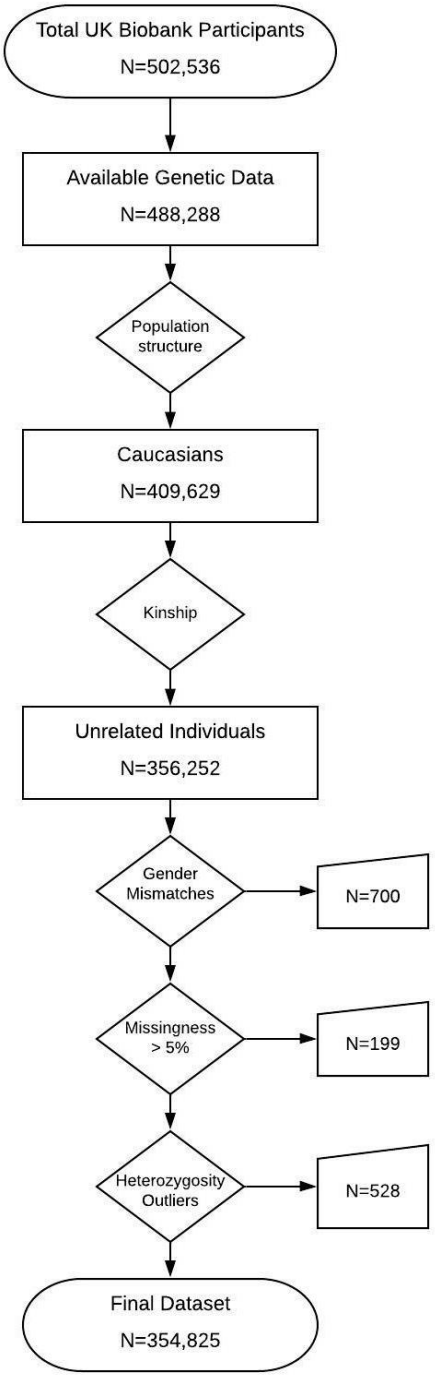


Figure 3.6 Flowchart indicating the selection of individuals for whom genetic data was available, and met the QC criteria.

A total of 147,711 participants were removed from the released dataset after filtering for ethnicity, kinship, gender mismatches, missingness and heterozygosity.

Table 3.3 Demographic characteristics of the UK Biobank cohort.

UK Biobank Demographics n=354,825	
Age at recruitment, years, mean (SD)	56.96 (7.91)
Male sex, n (%)	165,379 (46.61)
BMI, kg/m2, mean (SD)	27.42 (4.76)
Waist circumference, cm, mean (SD)	90.44 (13.49)
Alcohol intake, n (%)*	
Never	23,062 (6.50)
Special occasions	37,104 (10.46)
<1/week	39,064 (11.01)
1-2/week	93,237 (26.28)
3-4/week	85,926 (24.22)
Daily or almost daily	76,178 (21.47)
Smoking status, n (%)*	
Non-smoker	192,988 (54.39)
Ex-smoker	124,915 (35.20)
Current smoker	35,702 (10.06)
Comorbidities, n (%)	
Diabetes Mellitus	14,825 (4.18)
Hypertension	90,012 (26.64)
Hypercholesterolemia	41,535 (12.29)
Ischaemic Heart Disease	15,253 (4.51)
Cardiac failure	235 (0.07)
Chronic Kidney Disease categories*	
G1 (>90 ml/min	197,434 (55.64)
G2 (60-90 ml/min)	132,468 (37.33)
G3a (45-59 ml/min)	6,409 (1.81)
G3b (30-44 ml/min)	992 (0.28)
G4 (15-29 ml/min)	239 (0.07)
G5 (<15 ml/min)	89 (0.03)

*The following data were missing: alcohol intake for 0.07%, smoking satus for 0.34%, and CKD information for 4.85%. Diabetes, hypertension, hypercholesterolemia,ischaemic heart disease and cardiac failure were defined as present if they were self-reported as diagnosed by a doctor. Chronic Kidney Disease stages were defined as per the National Institute of Heath and Care Excellence (NICE) guidelines CG182.(NICE, 2015)

Table 3.4 Demographic characteristics of the PRS target dataset.

	All	PRS Target Dataset	
		Gout cases n=7,448	Controls n=39,959
Age at recruitment, years, mean (SD)	57.46 (7.94)	60.11 (6.86)	57.46 (7.94)
Male sex, n (%)	42,572 (89.80)	6,701 (89.97)	35,871 (89.77)
BMI, kg/m2, mean (SD)	28.11 (4.53)	30.64 (4.98)	27.63 (4.28)
SU, mg/dL, mean (SD)	5.92 (1.37)	6.65 (1.77)	5.78 (1.23)
Alcohol intake, n (%)*			
Never	2,441 (5.15)	355 (4.77)	2,086 (5.22)
Special occasions	3,410 (7.19)	455 (6.11)	2,955 (7.40)
<1/week	4,151 (8.76)	442 (5.93)	3,709 (9.28)
1-2/week	12,091 (25.50)	1,683 (22.60)	10,408 (26.05)
3-4/week	12,615 (26.61)	2,012 (27.01)	10,603 (26.53)
Daily or almost daily	12,660 (26.70)	2,489 (33.42)	10,171 (25.45)
Smoking status, n (%)*			
Non-smoker	23,152 (48.84)	3,040 (40.82)	20,112 (50.33)
Ex-smoker	18,849 (39.76)	3,711 (49.83)	15,138 (37.88)
Current smoker	5,232 (11.04)	668 (8.97)	4,564 (11.42)
Comorbidities, n (%)*			
Diabetes Mellitus	2,901 (6.12)	888 (11.92)	2,013 (5.04)
Hypertension	15,690 (33.10)	4,241 (56.94)	11,449 (28.65)
Hypercholesterolemia	7,655 (16.15)	2,031 (27.27)	5,624 (14.07)
Ischaemic Heart Disease	3,466 (7.31)	1,002 (13.45)	2,464 (6.17)
Cardiac failure	79 (0.17)	46 (0.62)	33 (0.08)
Chronic Kidney Disease categories*			
G1 (>90 ml/min	26,335 (55.55)	3,034 (40.74)	23,301(58.31)
G2 (60-90 ml/min)	19,472 (41.07)	3,629 (48.71)	15,843 (39.65)
G3a (45-59 ml/min)	1,177 (2.48)	515 (6.91)	662 (1.66)
G3b (30-44 ml/min)	280 (0.59)	181 (2.43)	99 (0.25)
G4 (15-29 ml/min)	82 (0.17)	63 (0.85)	19 (0.05)
G5 (<15 ml/min)	38 (0.08)	26 (0.35)	12 (0.03)

*The following data were missing: alcohol intake for 0.08%, smoking satus for 0.37%, and CKD information for 0.75%. Diabetes, hypertension, hypercholesterolemia,ischaemic heart disease and cardiac failure were defined as present if they were self-reported as diagnosed by a doctor. Chronic Kidney Disease stages were defined as per the National Institute of Heath and Care Excellence (NICE) guidelines CG182.(NICE, 2015)

3.3.3 PRS of gout vs. controls

A PRS for gout cases and controls was generated using PRSice, with the GUGC gout GWAS summary statistics as the base dataset, and the UK Biobank genotype data as the target dataset. The base dataset included 2,146,033 variants, compared to 717,091 in the target dataset; 209,025 variants were found in common among both datasets. After applying the clumping parameters, 90,188 remained available to generate the PRS for all cases and controls. The best-fit model to distinguish between gout cases and controls, was defined by the highest Nagelkerke's R^2. Table 3.5 provides a list of different p-value thresholds and the number of SNPs that contribute to each model. The greatest Nagelkerke's R^2 was 0.043, obtained with a p-value threshold of $7.0x10^{-6}$ (Figure 3.7). The best-fit model was calculated from association data of 10 SNPs, including ABCG2, SLC2A9, SLC17A3, SLC22A11 and CLNK loci (Table 3.6).

Table 3.5 PRSice results for a range of p-value thresholds and the number of SNPs that contributed to each model. The best-fit model was obtained at a p-value of $7.0x10^{-6}$ with a Nagelkerke's R^2 of 0.043.

p-value threshold	Nagelkerke's R2	p-value	# SNPs in the model
1	$9.63x10^{-5}$	0.101	90188
0.1	$2.2x10^{-4}$	0.012	15224
0.01	$8.8x10^{-4}$	$7.54x10^{-7}$	2003
0.001	$5.3x10^{-3}$	$2.83x10^{-34}$	267
0.0001	0.02	$1.02x10^{-122}$	43
0.00001	0.041	$2.24x10^{-235}$	13
$8.0x10^{-6}$	0.042	$3.10x10^{-238}$	12
$7.0x10^{-6}$	**0.043**	**$3.69x10^{-247}$**	**10**
$6.0x10^{-6}$	0.037	$3.40x10^{-216}$	8
$3.0x10^{-6}$	0.041	$6.55x10^{-231}$	6

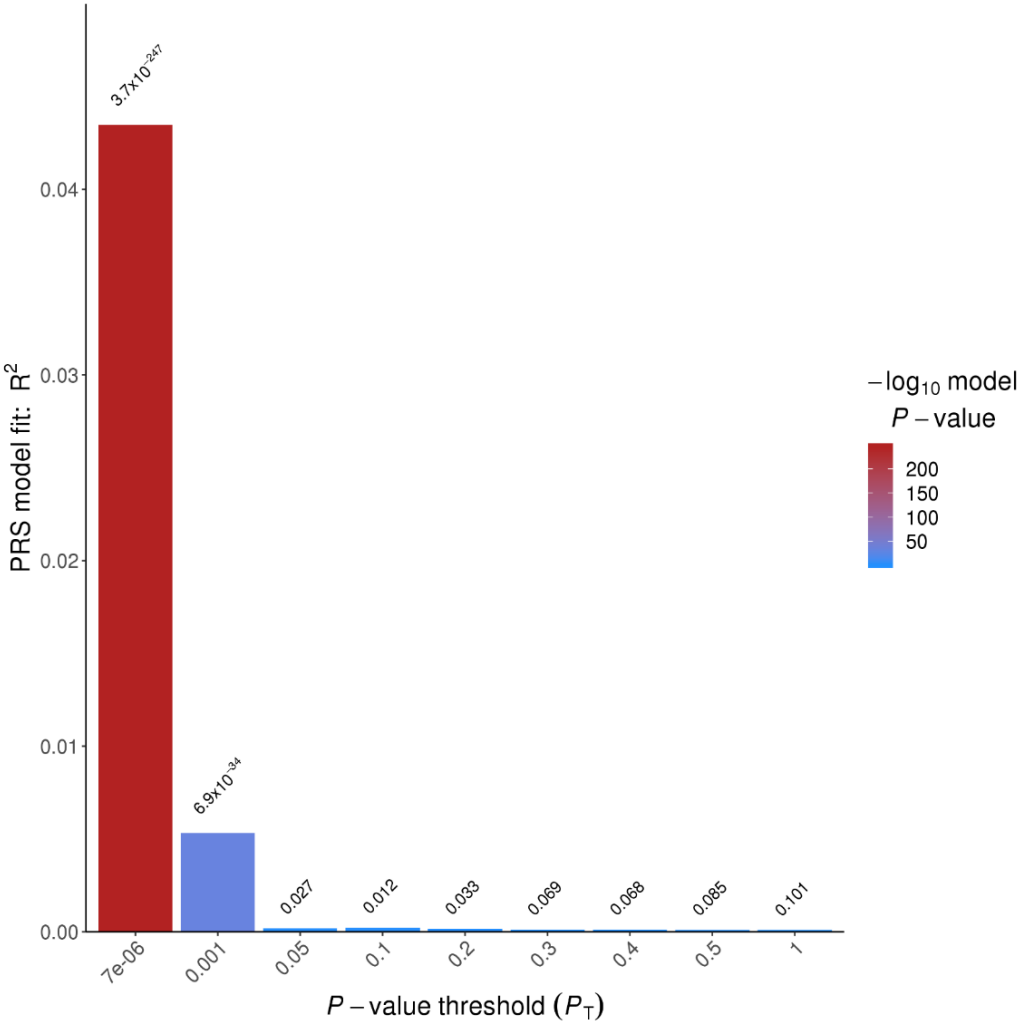


Figure 3.7 PRSice bar plot indicating the PRS model fit across different p value thresholds. The highest Nagelkerke's R^2 (0.043) was obtained with a p value=$7.0x10^{-6}$.

Table 3.6 SNPs under the best-fit p value threshold included in the PRS model for gout cases vs controls.

SNP	Chr	bp	Gene	P value
rs7442295	4	9966380	SLC2A9	$7.18x10^{-25}$
rs4698036	4	10331294	Intergenic	$5.56x10^{-18}$
rs997219	4	10524671	CLNK	$1.10x10^{-07}$
rs2231142	4	89052323	ABCG2	$4.98x10^{-32}$
rs2622604	4	89078924	ABCG2	$1.69x10^{-06}$
rs16891234	4	9946163	SLC2A9	$2.48x10^{-06}$
rs8087353	18	70669839	Intergenic	$4.99x10^{-06}$
rs7017745	8	118311980	Intergenic	$5.19x10^{-06}$
rs9393672	6	25842605	SLC17A3	$6.29x10^{-06}$
rs2078267	11	64334114	SLC22A11	$6.02x10^{-06}$

The mean (±SD) PRS for gout cases was 0.016 (±0.031) compared to 0.0019 (±0.032) for controls; this difference was significant using an independent t-test (p-value <0.0001). PRS were divided into deciles to explore the proportion of cases and controls. We observed an increase in gout prevalence across PRS deciles, going from 15% at the 2nd decile to an 82% at the 10th decile. Figure 3.8 shows the distribution of the scores for cases and controls and the proportion of cases and controls at each PRS decile. The predictive ability was assessed with the AUROC curve, which gave a 62% for the PRS model alone. The demographics model gave a 73% predictive ability, and included age, sex, BMI and SU into de regression model. The combination of the demographic characteristics and the PRS generated by the 10 SNPs generated a predictive ability of 75%. Figure 3.9 displays a comparison of the AUROC curves for the three models.

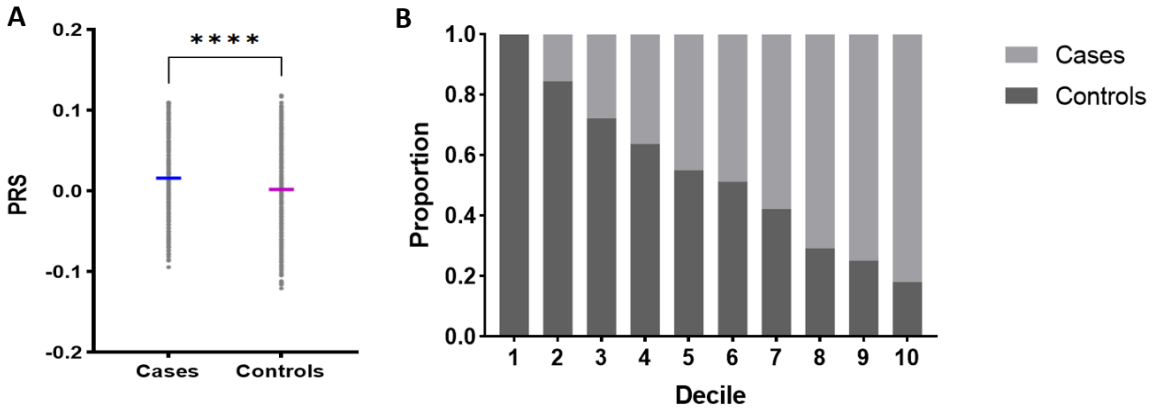
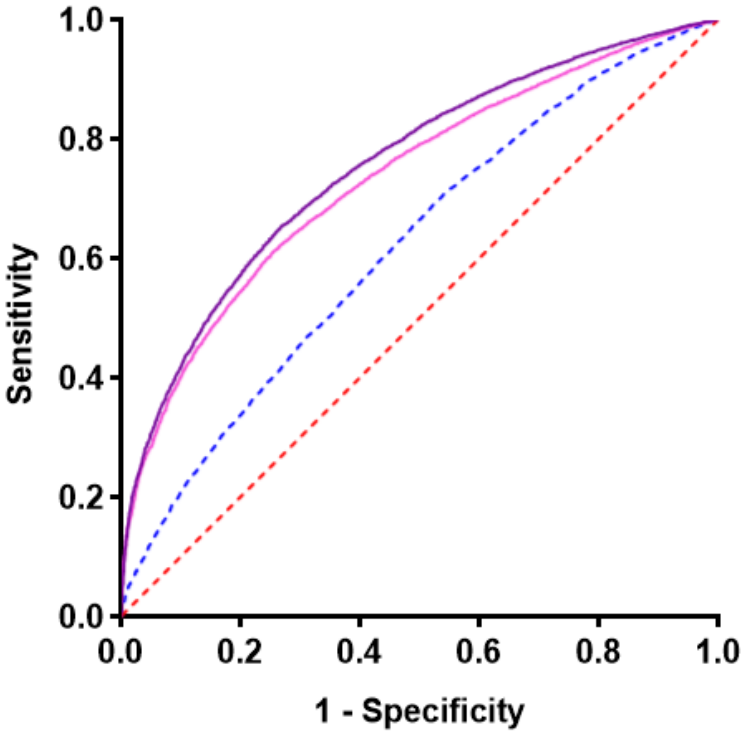


Figure 3.8 Distribution of polygenic risk scores (PRS) among gout cases and controls.

A) The horizontal lines represent the mean scores (0.016 for cases and 0.0019 for controls). The independent t-test showed a significant difference among groups (p <0.0001). B) Proportion of cases and controls at each decile of PRS.



Model	AUC
PRS + Covariates	0.75
Covariates (Age, Sex, BMI, SU)	0.73
PRS	0.62

Figure 3.9 Area under the receiver operating characteristics (AUROC) curve for the PRS model for gout vs. controls, compared to the demographics model, and combined model.

3.4 Discussion

This chapter described the quality control checks conducted on the UK Biobank genotype data. Although the UK biobank performed an internal QC evaluation, only few samples and markers were removed from the released dataset and additional variables were provided to filter information according research interests. For this thesis, association studies and PRS calculations required a more stringent revision to control for population structure, relatedness, missingness, heterozygosity and variants deviating from HWE. These criteria allowed the exclusion of genotyping errors, samples mishandling and systematic biases. Designing and applying robust QC procedures is crucial when conducting genetic association studies, as poor quality data would compromise the reliability of associations, and in consequence the results of downstream analyses. This QC pipeline was used to examine the datasets designed for the PRS for gout, described in this chapter, and for the GWAS and PRS for gout vs. asymptomatic hyperuricaemia, described in chapter 4.

PRS use the cumulative effects of genetic variants from GWAS summary statistics, and generate a score for each individual in an independent dataset (Euesden et al., 2014). The PRS model for gout cases vs. controls described here, was generated using the GUGC gout GWAS summary statistics published by Köttgen *et al* (Kottgen et al., 2013) as the training dataset, and the UK Biobank dataset containing genotype data for gout cases and controls as the test dataset. Köttgen's group published another GWAS for SU variation and gout risk (Adrienne Tin et al., 2019) that included >8 million SNPs for >400,000 individuals (including 13, 179 gout cases); and although using these GWAS summary statistics would have increased the SNPs in common between the base and test datasets, Tin's GWAS included genetic data of several UK

Biobank participants. Therefore, an alternative training dataset, though smaller, was used to avoid samples overlapping which can result in a significant inflation of the association results between the PRS and the phenotype.

PRS have the potential of identifying variants, that although not GWAS significant, when added up to other variants with small effects do contribute or are associated with a specific phenotype (S. M. Purcell et al., 2009). The PRS generated by PRSice in this study, showed a significant association with gout and were generated from 10 SNPs. Among these variants were the well-known loci located in the urate transporters ABCG2 and SLC2A9 genes (C.-J. Chen et al., 2018; Vitart et al., 2008), other variants located in or near SLC17A3, SLC22A11, CLNK and other intergenic regions. SLC22A11 and SLC17A3 are sodium phosphate transporters, and variants located in these genes have been associated to lower levels of serum urate and gout susceptibility, respectively (Dehghan et al., 2008; Riches, Wright, & Ralston, 2009; Masayuki Sakiyama, Matsuo, Shimizu, et al., 2016; Masayuki Sakiyama et al., 2014). CLNK is a cytokine-dependent cell linker that function as a positive regulator of immunoreceptor signalling (J. Yu et al., 2001). Several candidate gene association studies and whole genome sequencing data have shown that SNPs located in CLNK are associated with gout occurrence, with ORs ranging from 1.36 to 2.22 in Han Chinese population (T.-b. Jin et al., 2015), and OR=1.37 in Polynesian population (Ji et al., 2021). It has been hypothesised that CLNK contribute to gout risk by modulating the response upon cytokine stimulation via the STAT signalling pathway (Yao Zhang et al., 2016), confirming the importance of immune response in gout pathogenesis.

When the spread of PRS was compared among groups, the mean of gout cases was significantly higher than the mean PRS of controls. However, several

controls also had high PRS, in fact, the individual with the highest PRS was a control. This can indicate that although having risk loci, protective variants may be present as well, which reduces the risk of developing gout. Low PRS were also present in gout cases, which might indicate that even if genetic risk variants were not present, external factors (e.g. lifestyle factors) might be conditioning gout risk. These observations highlight the importance of performing more extensive genetic analyses to address the missing heritability, and of exploring additional environmental factors affecting the risk of gout (Chaudhury et al., 2018).

The predictive ability of the best-fit PRS model alone was 62%, compared to a 73% for demographic characteristics and 75% for the combined model. Previous studies have reported the use of polygenic risk scores to predict gout; however, their method differs from ours because they have used the association results of GWAS hits for serum urate variation. For instance, the largest GWAS published in 2019 by Tin *et al* conducted a PRS using the 187 SNPs that showed a significant association with SU to calculate risk for gout in UK Biobank participants. Their results showed an AUROC of 67% for genetic information alone and 84% when combined with demographic data (Adrienne Tin et al., 2019). This has been validated by Zhang & Lee in an independent cohort (MyCode Cohort) using 110 variants associated with SU, observing an AUROC of 64% for the PRS model and 80% for the demographics and PRS combined model (Yanfei Zhang & Lee, 2021). Another study has used PRS to determine their association with gout age of onset, presence of tophaceous gout and number of flares. In this study, they conducted a GWAS for gout using all the participants from the UK Biobank, and used the 12 independent SNPs that reached genome-wide significance to construct a genetic risk score for each individual of three independent cohorts. They observed a significant association

with gout (OR=1.21) and a negative association with age of onset (β=0.43) (Sumpter et al., 2019).

Although the predictive ability of the combined model generated in this project was lower than those generated in previous studies, ours was based only on the effect of 10 SNPs. Further research is required to determine whether this can be improved by using imputed data, which would certainly increase the number of common SNPs between the base and the test datasets. Several studies have shown that the predictive power can be enhanced by including more SNPs into the PRS, beyond those that reached genome-wide significance (Mavaddat et al., 2019; S. M. Purcell et al., 2009). On the other hand, this study was limited in terms of the base dataset used to generate the model, which had a lower number of gout cases. It would be interesting to evaluate if, when available, using larger summary statistics improve the PRS model.

PRS approaches have demonstrated to be suitable for various applications. First, in determining shared genetic components among traits; for instance among schizophrenia and bipolar disorder (S. M. Purcell et al., 2009), or among sporadic early-onset Alzheimer's disease (AD) and late-onset AD (Chaudhury et al., 2018). Second, in identifying individuals at higher risk of developing complex diseases. This is expected to benefit from a preventive medicine perspective, by initiating environmental or lifestyle modifications that are important in a disease progression, but also for the selection of participants for clinical trials (Hu et al., 2013). Introducing preventive measures is critical in gout management, and people at higher risk of developing the disease or presenting a more severe form, could benefit from early interventions such as lifestyle changes. Additionally, there are no clinical trials that have looked at the impact of long-term treat-to-target strategy with ULT, and this has become a necessity

in gout research (Perez-Ruiz & Dalbeth, 2019). In this regard, PRS models could play an important role in the design of the trial.

This study presented several limitations. First, gout was only defined via self-reports, ULT use and hospital records, rather than using the ACR/EULAR classification criteria. Second, age is a well known confounder for gout, however we only matched cases and controls by gender. This was done in order to maximise the number of controls in our datasets, and we controlled for the confounding effect by adding age as a covariate in the demographics and combined models. However, we acknowledge that matching by age and gender could have been more appropriate to determine their effect on gout risk. Additionally, data on medication were not included as covariates in the PRS model. It is well known that several diuretics alter SU levels and may contribute to an increased risk of gout. This is suggested for future studies to determine whether the inclusion of medication and comorbidity data improves the predictive ability of the demographics model. Finally, the test dataset was comprised by direct genotyping data instead of imputed data. While this represents a computationally intensive process, it might well improve the predictive power and is suggested as future research.

In conclusion, this chapter stablished the QC pipeline to clean genotype data for GWAS and PRS analyses. The PRS model for gout was generated from 10 SNPs, that included urate transporters and one gene involved in the immune response. Validation in independent datasets is required.

Chapter 4. Genome Wide Association Study and Polygenic Risk Score of Asymptomatic Hyperuricaemia vs. Gout.

4.1 Introduction

Gout is a common form of inflammatory arthritis caused by the deposition of monosodium urate (MSU) crystals. Elevated serum urate (SU) concentration is the precursor to MSU crystal deposition, and the onset of gout (Nicola Dalbeth et al., 2016). However, the majority of people with hyperuricemia do not develop gout. For instance, in the USA, the prevalence of hyperuricaemia (defined as SU >7.0 mg/dL) is 20%, while that of gout is 3.9% (Chen-Xu, Yokose, Rai, Pillinger, & Choi, 2019). The reason(s) why only some people with hyperuricaemia develop gout is poorly understood. Genome wide association studies (GWAS) have improved the understanding of the pathophysiology of hyperuricaemia and gout over the last 10-15 years. For instance, genetic variants located in urate transporters such as the ABCG2, SLC2A9, and SLC22A11 genes have been identified as risk loci for both hyperuricemia and gout (Dehghan et al., 2008; Kolz et al., 2009; Kottgen et al., 2013; S. Li et al., 2007). Additional genetic variants such as GCKR, PDZK1 and ALDH2 that play important roles in glucose, cholesterol and alcohol metabolism, respectively, have been associated with both phenotypes (Kottgen et al., 2013; C. Li et al., 2015; H. Matsuo et al., 2016; Phipps-Green et al., 2016). However, the genetic variants associated with progression from hyperuricaemia to gout remain poorly understood. To date, only a single GWAS in 6,009 Japanese adults (2,860 with gout) has examined this, and revealed two novel loci: CNTN5 and MIR302F,

that participate in immune and inflammatory responses (Kawamura et al., 2019). Further analyses in independent populations and larger sample sizes are needed to improve the understanding of the molecular mechanisms involved in transitioning from asymptomatic hyperuricaemia to gout.

Thus, the purpose of this study was to examine the genetic variants associated with transition from asymptomatic hyperuricaemia to gout. In order to meet this objective, a GWAS using gout cases and asymptomatic hyperuricaemia controls (defined as SU ≥6.0 mg/dL) derived from the UK Biobank resource was performed. A threshold of ≥6.0 mg/dL to define AH was chose, as the risk of incident gout increases above this SU level (N. Dalbeth et al., 2018). Genotype data was used to develop a polygenic risk score (PRS) to predict gout-case and asymptomatic hyperuricaemia-control status. Finally, in order to compare the effect of the independent SNPs from the gout vs. asymptomatic hyperuricaemia GWAS, two additional GWAS were performed using gout cases vs. normouricaemia controls with SU <6.0 mg/dL, and gout vs. normouricaemia controls with SU <7.0 mg/dL.

4.2 Materials and Methods

4.2.1 UK Biobank cohort

This study was conducted using data from the UK Biobank resource (Project ID 45987). Details about the UK Biobank, the samples processing, genotyping arrays and internal quality control criteria have been described in Chapter 3, sections 3.2.1-3.2.4. Briefly, the UK Biobank is a prospective study of ~500,000 participants, aged 40-60 years and recruited across England, Wales and Scotland between the years 2006 and 2010. Data were collected on lifestyle and sociodemographic information, cognitive function, health status, and family medical history. Participants had standard physical and functional measurements, and provided blood samples for genetic analyses. Details about recruitment and samples processing for genotyping are described elsewhere (Sudlow et al., 2015; Welsh, 2017).

4.2.2 Subjects

This study included the generation of a GWAS and a PRS of gout cases vs asymptomatic hyperuricaemia controls, and a GWAS of gout cases vs normouricaemia controls (Figure 4.1). Therefore, three different phenotypes were derived from the UK Biobank cohort. The phenotypes were defined as follows:

Gout cases: Gout was defined as present if any of the following criteria were met: self-reported physician diagnosed gout; urate lowering therapy (ULT) prescription without a hospital diagnosis of lymphoma or leukaemia (ICD-10 codes C81-C96); or a primary or secondary diagnosis of gout in hospital discharge letters using the ICD-10 codes M10, M100-M14, and M109. Participants with self-reported physician

diagnosed gout were excluded if their SU was <6.0 mg/dL and they did not report prescription of ULT at the UK Biobank visit.

Asymptomatic hyperuricaemia controls: Participants with SU ≥6.0 mg/dL and not classified as gout. Hyperuricaemia is usually defined as SU ≥6.8 mg/dL, which is the supersaturation concentration under physiological conditions (Loeb, 1972), or SU ≥7.0 mg/dL as per the National Health and Nutrition Examination Survey (PPC) definition (Zhu et al., 2011). However, for this study, a threshold of ≥6.0 mg/dl was chosen as it associates with incident gout in prospective studies (N. Dalbeth et al., 2018).

Normouricaemia controls: Participants with SU <6.0 mg/dL and not classified as gout were considered as normouricaemia controls. Given the uncertainty around definition of normal SU, for example SU <6.0 mg/dL being the treatment threshold for treat-to-target ULT (Ruoff & Edwards, 2016), while epidemiological studies use a cut-off of <7.0 mg/dl (Zhu et al., 2011), another group of normouricaemia controls was ascertained with SU <7.0 mg/dl and not classified as gout.

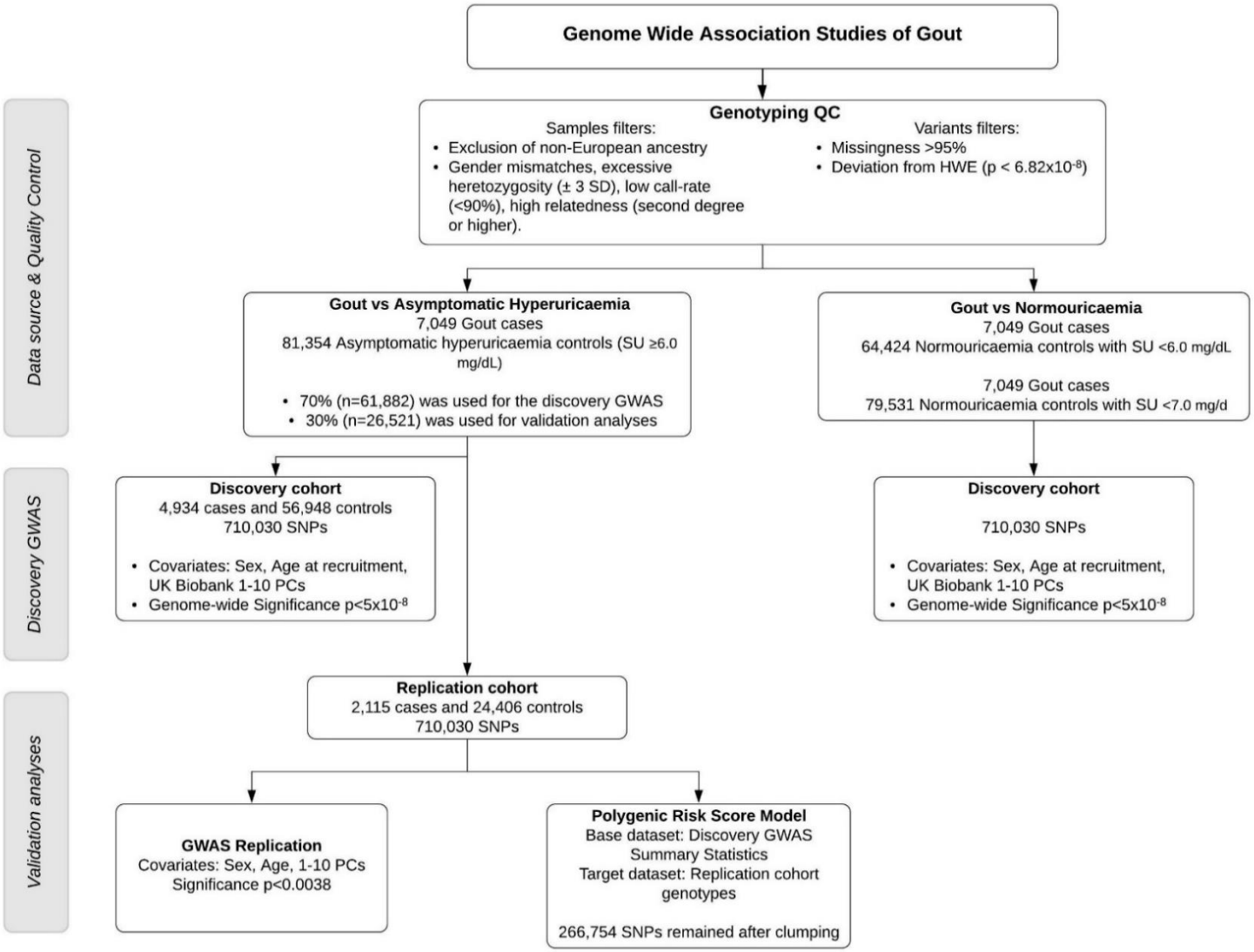


Figure 4.1 Study design. Workflow for the discovery and replication analyses.

QC, quality control; SD, standard deviation; HWE, Hardy-Weinberg equilibrium; SNPs, single nucleotide polymorphisms; PCs, principal component

4.2.3 Genotyping quality control

The UK Biobank centrally performed QC procedures have been described in Chapter 3 section 3.2.4. After excluding the individuals whose genotyping data did not meet the QC criteria (i.e. non-European ancestry, gender mismatches, kinship, and missingness and heterozygosity extreme outliers), 354,825 participants comprised the final working dataset. For this study, two cohorts were derived: Gout vs. asymptomatic hyperuricaemia, and gout vs. normouricaemia. In order to reduce the computing time when conducting further stringent QC filters, both cohorts were generated by extracting the individuals of interest and assigning the phenotype that corresponded. Additional QC refinement was conducted using the command line software PLINK version 1.9 (Chang et al., 2015; S. Purcell et al., 2007). All the commands used for each parameter are detailed in Chapter 3 section 3.2.5. Briefly, individuals with a kinship coefficient equivalent to second degree (or greater) relatives were excluded. Individuals were also excluded if they had a heterozygosity ±3 standard deviations (SD) from the mean, or a call rate <90%. Markers with a call rate <95%, or those deviating from Hardy-Weinberg equilibrium (Bonferroni corrected p value threshold $p=6.82x10^{-8}$) were removed from the dataset.

4.2.4 Genome-wide association studies

Gout vs. AH GWAS: This cohort was divided into two datasets: 70% (n=61,882) was used as the discovery dataset, and the remaining 30% (n=26,521) was used as the validation dataset for replication analysis and PRS generation (Figure 4.1).

Gout vs. normouricaemia GWAS: For this cohort, two separate GWAS were conducted, using gout cases vs. 64,424 controls with SU <6.0 mg/dL, and gout cases vs. 79,531 controls with SU <7.0 mg/dL.

Discovery and replication association tests were performed using PLINK v1.9. Odds ratios (OR) and 95% confidence intervals (CI) were computed using additive logistic regression with the following command:

```
>plink --bfile INPUT_BINARY-FILE_NAME --logistic --ci 0.95
--adjust --covar COVARIATES_FILE.txt --out OUTPUT_FILE_NAME
```

Binary files containing individuals' information, genotype data and genetic markers information (.fam, .bed, .bim files, respectively) for each cohort, were used as the input file. Given that this study involved the analysis of genetic variants and their association with a binary outcome (i.e. gout cases vs. asymptomatic hyperuricaemia controls, and gout cases vs. normouricaemia controls), logistic regression was performed, with the inclusion of age when attended the baseline visit, sex and the first 10 principal components (PCs) as covariates. To control for multiple testing, a genome-wide significance threshold of 5.0×10^{-8} was used. Figure 4.2 provides an overview of the input files and covariates file content.

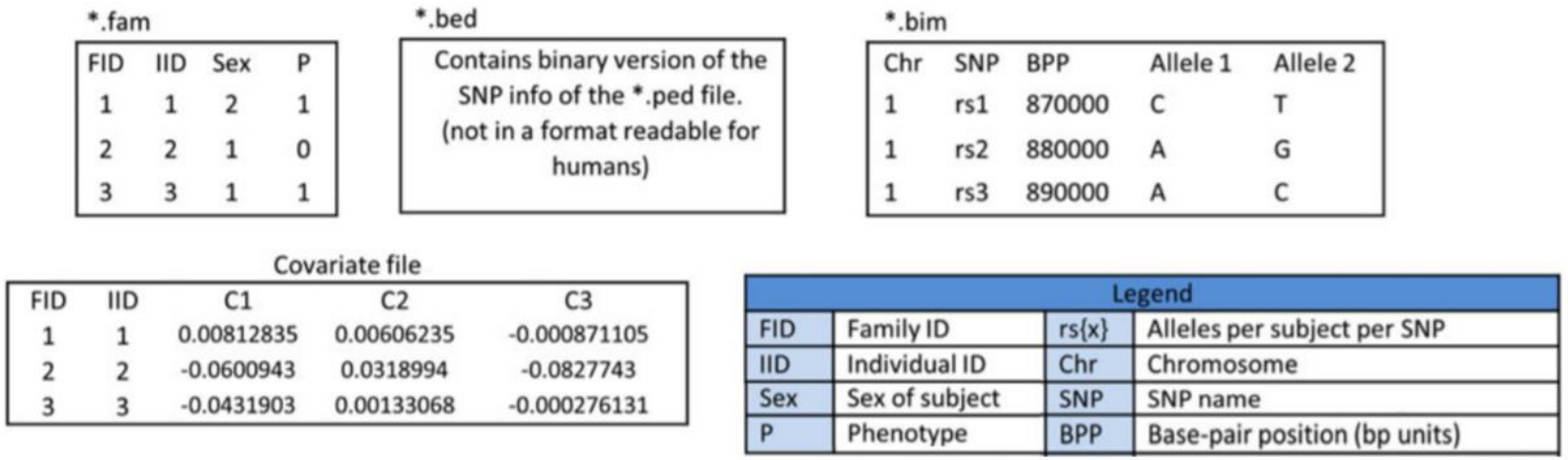


Figure 4.2 Overview of PLINK binary files format and covariates file.
Adapted from (Marees et al., 2018).

Genetic variants might be correlated due to linkage disequilibrium (LD); thus, despite reaching genome-wide significance, they would not be independent. Therefore, to determine the number of independent loci from the GWAS analysis, LD clumping was performed using the following command, where SNPs with a p value $<1x10^{-5}$, r^2 >0.1, and a 500kb window from the index SNPs were assigned to the clump:

```
>plink --bfile INPUT_BINARY-FILE_NAME --clump SUMMARY-
STATS.assoc.logistic --clump-p1 0.00000001 --clump-p2 0.00001
--clump-r2 0.1 --clump-kb 500 --out OUTPUT_FILE_NAME
```

Additionally, pairwise LD was further analysed using the R package LDlinkR (Myers, Chanock, & Machiela, 2020). This package provides access to the web-based tool LDlink, which is designed to explore variation and LD structure across 26 population groups, using the Phase 3 release of the 1000 Genomes Project (1000G) as the reference panel. LDlink includes multiple applications, presented in eight different modules: LDassoc, LDexpress, LDhap, LDmatrix, LDpair, LDpop, LDproxy and LDtrait. LDmatrix is the module that allows the generation of heat maps of pairwise comparisons, based on LD metrics (i.e. D^{l} and R^2) given a list of SNPs with RS identifiers and a population of interest (Machiela & Chanock, 2015). For this study, the SNPs identified as independent located in the same gene or less than 500kb apart were included in the pairwise analysis. Finally, HaploView was used to generate the LD plot based on R^2 values (Barrett, Fry, Maller, & Daly, 2004).

4.2.5 Annotation

Annotation of GWAS results was conducted using the Functional Mapping and Annotation of Genome Wide Association Studies (FUMA-GWAS). FUMA is a web-based post-GWAS tool that integrates different repositories including LD patterns and functional data of a wide variety of SNPs, through two separate processes: SNP2GENE and GENE2FUNC. For this project, the SNP2GENE tool was used to confirm independent and significant SNPs, and to map them to genes and to generate the Manhattan plots. SNP2GENE requires the summary statistics in PLINK format as the input file, it then identifies independent SNPs based on their LD structure and the parameters selected for the analysis, and independent SNPs are finally mapped to genes according to their functional consequences (i.e. physical position, expression quantitative trait loci and chromatin interactions). Table 4.1 describes the selected parameters for the identification of lead SNPs using the SNP2GENE tool. (Watanabe, Taskesen, van Bochoven, & Posthuma, 2017)

Table 4.1 FUMA-SNP2GENE parameters for lead SNPs identification

Parameter	Value
Sample size	N[1]
Maximum p value of lead SNPs	5.0×10^{-8}
Maximum p value cut-off	0.05
R2 threshold to define independent significant SNPs	0.1
Reference panel population	1000G Phase 3 (EUR)[2]
Minimum Minor Allele Frequency	0.01
Maximum distance between LD blocks to merge into a locus	500kb

[1]61882 for gout vs Asymptomatic hyperuricaemia GWAS, 85580 for gout vs SU <7.0 mg/dL, and 71,473 for gout vs SU <6.0 mg/dL. [2]Phase 3 release of the 1000 Genomes Project, European population.

4.2.6 Replication analysis: gout cases vs. asymptomatic hyperuricaemia controls

For replication analysis, the variants that reached genome wide significance in the discovery analysis were tested for association with gout in the replication cohort. Logistic regression was adjusted for sex, age, and the first 10 PCs. A Bonferroni corrected p value of <0.004 (0.05/13) was used to determine significant associations in the replication analysis. The results from the discovery GWAS and the replication analysis were combined by meta-analysis using PLINK. The fixed-effects model was used to estimate pooled ORs and 95% CI, and Cochran's Q test *p* values and I^2 values were used to assess heterogeneity.

4.2.7 Additional statistical analyses

Baseline data were summarised using mean and standard deviation (SD) for continuous variables, and number (%) for categorical variables. Independent sample t-test and chi-square test were used to compare continuous and categorical data respectively.

As previous GWAS (Kawamura et al., 2019; Nakayama et al., 2017) have used a cut-off of 7.0 mg/dl to define hyperuricaemia, a sensitivity analysis was conducted to evaluate if the association of GWAS hits and gout remained significant if controls had a SU ≥7.0 mg/dL.

Linear regression was used to examine the effect of GWAS hits on SU levels. This was performed using the full cohort, and adjusted for sex and age at recruitment. Beta coefficients and standard errors (SE), and adjusted Beta coefficients and SE were calculated.

4.2.8 Polygenic risk scoring of gout vs. asymptomatic hyperuricaemia

PRS was calculated using PRSice-2 (S. W. Choi & O'Reilly, 2019). The discovery GWAS summary statistics were used as the base dataset, while the replication cohort genotype-phenotype data were used as the target dataset. Clumping parameters in PRSice were set to an r^2 >0.1 and a 500kb window from the index SNPs. The lower and upper p value thresholds were set as 0 and 1, respectively, in increments of $1.0x10^{-6}$. PRS was generated using the following commands and parameters (Table 4.2 includes a description of each command). Logistic regression was used to estimate the effect of age at recruitment, sex and BMI using SPSS Statistics 24. AUROC was used to evaluate the predictive ability of the best-fit PRS, the demographic characteristics and combined models.

```
Rscript PRSice.R --dir . \
--prsice ./PRSice \
--base DISCOVERY-SUMMARY-STATS.assoc.logistic \
--target REPLICATION-COHORT_BINARY-FILE_NAME \
--thread max \
--print-snp T \
--stat OR \
--binary-target T \
--no-clump F \
--clump-kb 500 \
--clump-r2 0.1 \
--clump-p 1 \
--perm 10000 \
--lower 0 \
--interval 0.000001 \
--upper 1 \
--out AH_gout_PRS_kb500
```

Table 4.2 PRSice commands used to generate the PRS.

Command	Description	Selected parameter
--base	Base/Training dataset	Discovery GWAS Summary statistics
--target	Target dataset	Replication cohort
--thread	Number of thread	Maximum
--print-snp	Generates an output with the SNPs that remain after clumping	TRUE
--stat	Column header containing the effect size estimate	Odds Ratios
--binary-target	Indicates if the target dataset includes a dichotomous phenotype	TRUE
--no-clump	If true, PRSice do not compute LD clumping	FALSE
--clump-kb	Indicates the window for clumping	500 kb
--clump-r2	Indicates the r2 threshold for clumping	0.1
--clump-p	Indicates the p value threshold for clumping	1
--perm	Number of permutation to generate the empirical p value	10000
--lower	Indicates the starting p value threshold to be tested	0
--interval	Indicates the step size of the p value threshold	0.000001
--upper	Indicates the final p value threshold to be tested	1
--out	Output name	FILE_NAME

4.3 Results

4.3.1 Demographic characteristics

Gout vs. Asymptomatic hyperuricaemia: Following genotype QC filters, data for 7,049 gout cases and 81,354 AH controls were included. The entire cohort was comprised of 80.77% men, and their mean (SD) age, BMI, and SU were 57.87 (7.77) years, 29.63 (4.81) kg/m^2, and 6.92 (0.88) mg/dL, respectively. This cohort was divided into the discovery and replication datasets (Table 4.3, Figure 4.1). The two datasets had comparable disease and demographic characteristics.

Gout vs. normouricaemia: Two separate GWAS were conducted using the 7,049 gout cases, comprised by 91.91% men. The mean (SD) age, BMI and SU were 60.09 (6.86) years, 30.74 (4.95) kg/m^2 and 6.75 (1.77) mg/dL, respectively. 79,531 controls with SU <7.0 mg/dL and 64,424 controls with SU <6.0 mg/dL were included in each GWAS. These groups were comprised of 41.85% and 33.94% men, and their mean (SD) age, BMI and SU were 56.75 (7.95) and 56.53 (7.99) years; 27.06 (4.62) and 26.61 (4.50) kg/m^2; and 4.92 (1.07) and 4.57 (0.86) mg/dL, respectively. Table 4.4 describes the demographics, lifestyle and comorbidities for gout cases and normouricaemia controls.

Table 4.3 Demographic, life-style and comorbidities for gout cases and asymptomatic hyperuricaemia controls of the UK Biobank

		Discovery GWAS		Replication stage	
	All	**Gout cases**	**Controls**	**Gout cases**	**Controls**
	n=88,403	n=4,934	n=56,948	n=2,115	n=24,406
Age at recruitment, years, mean (SD)	57.87 (7.77)	60.11 (6.88)	57.67 (7.82)	60.06 (6.81)	57.69 (7.80)
Male sex, n (%)	71,401 (80.77)	4,529 (91.79)	45,506 (79.91)	1,950 (92.20)	19,416 (79.55)
BMI, kg/m^2, mean (SD)	29.63 (4.81)	30.79 (4.96)	29.53 (4.78)	30.62 (4.93)	29.53 (4.78)
SU, mg/dL, mean (SD)	6.92 (0.88)	6.74 (1.78)	6.94 (0.76)	6.77 (1.74)	6.93 (0.75)
Alcohol intake, n (%)*					
Never	4,397 (4.97)	221 (4.48)	2,822 (4.96)	85 (4.02)	1,269 (5.20)
Special occasions	6,698 (7.58)	272 (5.51)	4,405 (7.74)	115 (5.44)	1,906 (7.81)
<1/week	7,588 (8.58)	278 (5.63)	5,006 (8.79)	128 (6.05)	2,176 (8.92)
1-2/week	21,908 (24.78)	1,093 (22.15)	14,190 (24.92)	496 (23.45)	6,129 (25.11)
3-4/week	23,370 (26.44)	1,351 (27.38)	15,035 (26.40)	589 (27.85)	6,395 (26.20)
Daily or almost daily	24,365 (27.56)	1,710 (34.66)	15,448 (27.13)	701 (33.14)	6,506 (26.66)
Smoking status, n (%)*					
Non-smoker	41,778 (47.26)	1,998 (40.49)	27,243 (47.84)	862 (40.76)	11,675 (47.84)
Ex-smoker	37,705 (42.65)	2,481 (50.28)	23,871 (41.92)	1,065 (50.35)	10,288 (42.15)
Current smoker	8,590 (9.72)	434 (8.80)	5,613 (9.86)	183 (8.65)	2,360 (9.67)
Comorbidities, n (%)					
Diabetes Mellitus	5,389 (6.10)	589 (11.94)	3,205 (5.63)	245 (11.58)	1,350 (5.53)
Hypertension	35,776 (40.47)	2,824 (57.24)	22,206 (38.99)	1,241 (58.68)	9,505 (38.95)
Hypercholesterolemia	15,831 (17.91)	1,368 (27.73)	9,737 (17.10)	585 (27.66)	4,141 (16.97)
Ischaemic Heart Disease	6,817 (7.71)	666 (13.50)	4,051 (7.11)	290 (13.71)	1,810 (7.42)
Cardiac failure	133 (0.15)	28 (0.57)	62 (0.11)	18 (0.85)	25 (0.10)
Chronic Kidney Disease stages*					
G1 (>90 ml/min	39,800 (45.02)	1,962 (39.76)	25,849(45.39)	855 (40.43)	11,134 (45.62)
G2 (60-90 ml/min)	43,774 (49.52)	2,434 (49.33)	28,260 (49.62)	1,026 (48.51)	12,054 (49.39)
G3a (45-59 ml/min)	3,722 (4.21)	343 (6.95)	2,258 (3.97)	159 (7.52)	962 (3.94)
G3b (30-44 ml/min)	792 (0.89)	138 (2.80)	419 (0.74)	40 (1.89)	195 (0.80)
G4 (15-29 ml/min)	213 (0.24)	41 (0.83)	106 (0.19)	22 (1.04)	44 (0.18)
G5 (<15 ml/min)	51 (0.06)	13 (0.26)	19 (0.03)	12 (0.57)	7 (0.03)

*The following data were missing: alcohol intake for 0.09%, smoking satus for 0.73%, and CKD information for 0.06%. Diabetes, hypertension, hypercholesterolemia,ischaemic heart disease and cardiac failure were defined as present if they were self-reported as diagnosed by a doctor. Chronic Kidney Disease stages were defined as per the National Institute of Heath and Care Excellence (NICE) guidelines CG182.(NICE, 2015).

Table 4.4 Demographic, life-style and comorbidities for gout cases and normouricaemia controls of the UK Biobank.

	Gout vs SU<7.0 mg/dL		Gout vs SU<6.0 mg/dL	
	Gout cases	**Controls**	**Gout cases**	**Controls**
	n=7,049	n=79,531	n=7,049	n=64,424
Age at recruitment, years, mean (SD)	60.09 (6.86)	56.75 (7.95)	60.09 (6.86)	56.53 (7.99)
Male sex, n (%)	6,479 (91.91)	33,280 (41.85)	6,479 (91.91)	21,867 (33.94)
BMI, kg/m2, mean (SD)	30.74 (4.95)	27.06 (4.62)	30.74 (4.95)	26.61 (4.50)
SU, mg/dL, mean (SD)	6.75 (1.77)	4.92 (1.07)	6.75 (1.77)	4.57 (0.86)
Alcohol intake, n (%)*				
Never	306 (4.35)	5,387 (6.78)	306 (4.35)	4,572 (7.10)
Special occasions	387 (5.50)	8,713 (10.96)	387 (5.50)	7,444 (11.56)
<1/week	406 (5.77)	8,968 (11.28)	406 (5.77)	7,588 (11.79)
1-2/week	1,589 (22.57)	21,254 (26.74)	1,589 (22.57)	17.311 (26.89)
3-4/week	1,940 (27.56)	19,156 (24.10)	1,940 (27.56)	15,183 (23.58)
Daily or almost daily	2,411 (34.25)	16,006 (20.14)	2,411 (34.25)	12,2856 (19.08)
Smoking status, n (%)*				
Non-smoker	2,860 (40.57)	44,226 (55.61)	2,860 (40.57)	36,718 (56.99)
Ex-smoker	3,546 (50.31)	27,057 (34.02)	3,546 (50.31)	21,036 (32.65)
Current smoker	617 (8.75)	7,995 (10.05)	617 (8.75)	6,465 (10.04)
Comorbidities, n (%)				
Diabetes Mellitus	834 (11.83)	2,962 (3.72)	834 (11.83)	2,235 (3.47)
Hypertension	4,065 (57.67)	18,778 (23.61)	4,065 (57.67)	13,644 (21.18)
Hypercholesterolemia	1,953 (27.71)	8,942 (11.24)	1,953 (27.71)	6,517 (10.12)
Ischaemic Heart Disease	956 (13.56)	3,107 (3.91)	956 (13.56)	2,158 (3.35)
Cardiac failure	46 (0.65)	32 (0.04)	46 (0.65)	18 (0.03)
Chronic Kidney Disease categories*				
G1 (>90 ml/min	2,798 (39.72)	48,410 (60.91)	2,798 (39.72)	41,026 (63.73)
G2 (60-90 ml/min)	3,477 (49.35)	29,868 (37.58)	3,477 (49.35)	22,617 (35.13)
G3a (45-59 ml/min)	502 (7.13)	1,084 (1.36)	502 (7.13)	683 (1.06)
G3b (30-44 ml/min)	180 (2.56)	83 (0.10)	180 (2.56)	39 (0.06)
G4 (15-29 ml/min)	63 (0.89)	20 (0.03)	63 (0.89)	8 (0.01)
G5 (<15 ml/min)	25 (0.35)	11 (0.01)	25 (0.35)	6 (0.01)

*The following data were missing: alcohol intake for 0.07%, smoking satus for 0.32%, and CKD information for 0.07%. Diabetes, hypertension, hypercholesterolemia,ischaemic heart disease and cardiac failure were defined as present if they were self-reported as diagnosed by a doctor. Chronic Kidney Disease stages were defined as per the National Institute of Heath and Care Excellence (NICE) guidelines CG182.(NICE, 2015)

4.3.2 GWAS: Gout vs. Asymptomatic Hyperuricaemia

An additive logistic regression was performed to test the association between gout and 710,030 genetic variants. Thirty-four SNPs reached genome wide significance ($p<5.0x10^{-8}$) and after filtering for LD ($r^2<0.1$), 13 SNPs were identified as independent associations. Figure 4.3 shows the Manhattan plot for this GWAS. These 13 lead SNPs were selected for the replication study, where they were tested for association with gout in the remaining 30% of the dataset. Successful replication was defined if the p value was < 0.004. Summary results for both the discovery and the replication analyses are shown in Table 4.5. The SNP with the greatest effect was rs2231142 in ABCG2 gene with aOR=1.66 ($2.05x10^{-78}$) in the discovery stage, and aOR=1.64 ($1.17x10^{-32}$) in the replication stage. This was followed by a novel locus: rs1229984 in ADH1B gene with aOR=1.51, $p=5.00x10^{-12}$; and aOR=1.44, $p=4.77x10^{-5}$ for the discovery and replication analysis, respectively. The remaining SNPs were located in or near GCKR, PPM1K-DT, SLC2A9, MEPE, LOC105377323, and SLC22A11. Although an $r^2<0.1$ was set as the threshold to determine independent associations in PLINK and FUMA, because eight SNPs were located in chromosome 4, near the ABCG2 locus, additional pairwise LD comparisons were conducted using LDlinkR and Haploview for those variants located <500kb apart (Figure 4.4). This analysis did not show evidence of tight LD, and the 13 SNPs were retained as independent associations for further analyses.

A sensitivity analysis was conducted to examine whether the association between the 13 lead SNPs and gout remained significant when controls with asymptomatic hyperuricaemia with SU <7.0 mg/dL were excluded. Logistic regression was

adjusted for age and sex, and corrected for multiple testing using Bonferroni. For all variants, the ORs diminished in magnitude but remained significant (Table 4.6).

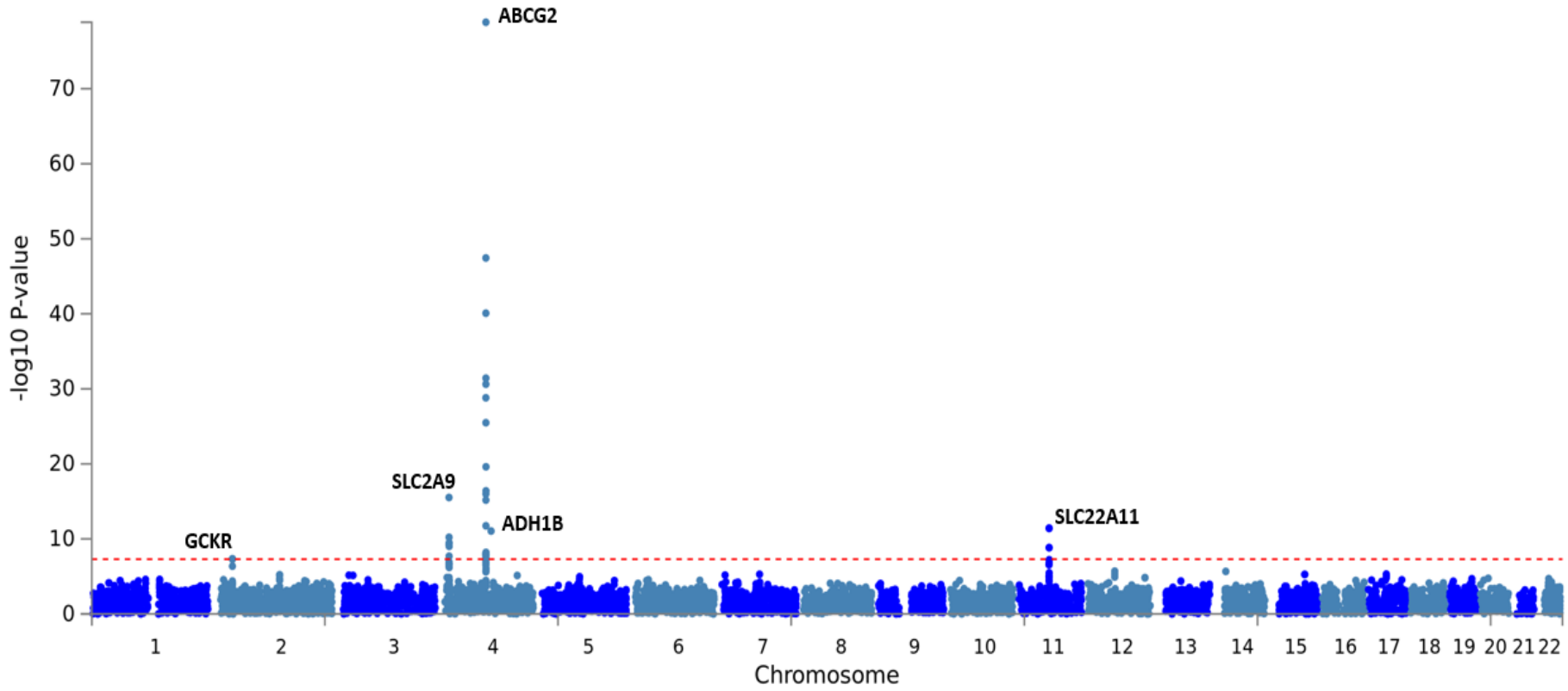


Figure 4.3 Manhattan plot of the discovery GWAS of gout vs. AH (SU ≥6.0 mg/dL) controls.

The y-axis shows –log10 P values ordered by chromosomal position on the x-axis. The horizontal dashed-line represents genome wide significance threshold (5.0×10^{-8}).

Table 4.5 Summary of GWAS and replication analysis of 13 lead SNPs in gout cases and AH controls (SU ≥6.0 mg/dL).

						Discovery GWAS		Replication stage		Meta-Analysis[1]		
SNP	Chr	bp	Gene	A1	Freq	aOR (95% CI)[2]	P value	aOR (95% CI)[2]	P value	aOR (95% CI)[2]	P value	Cochrane's Q p value
rs1260326	2	27730940	GCKR	T	0.41	1.13 (1.08-1.17)	$3.54x10^{-08}$	1.15 (1.08-1.23)	$2.54x10^{-05}$	1.14 (1.10-1.18)	$5.10x10^{-12}$	0.61
rs2231142	4	89052323	ABCG2	T	0.14	1.66 (1.58-1.76)	$2.05x10^{-78}$	1.64 (1.51-1.78)	$1.17x10^{-32}$	1.65 (1.58-1.73)	$3.33x10^{-109}$	0.75
rs13120400	4	89033527	ABCG2	C	0.28	0.82 (0.78-0.86)	$1.56x10^{-16}$	0.84 (0.78-0.91)	$3.43x10^{-06}$	0.83 (0.79-0.86)	$3.66x10^{-21}$	0.50
rs7672194	4	89126647	ABCG2	T	0.48	1.16 (1.11-1.21)	$3.21x10^{-12}$	1.15 (1.08-1.23)	$1.13x10^{-05}$	1.16 (1.12-1.20)	$1.58x10^{-16}$	0.88
rs4693211	4	89249061	PPM1K-DT	C	0.07	1.41 (1.31-1.52)	$6.97x10^{-20}$	1.26 (1.12-1.42)	$1.21x10^{-04}$	1.37 (1.28-1.45)	$1.55x10^{-22}$	0.11
rs28793136	4	89216768	PPM1K-DT	C	0.08	1.35 (1.26-1.45)	$8.19x10^{-17}$	1.26 (1.12-1.40)	$6.29x10^{-05}$	1.32 (1.25-1.40)	$4.79x10^{-20}$	0.27
rs1545207	4	89239492	PPM1K-DT	A	0.28	1.14 (1.09-1.20)	$6.82x10^{-09}$	1.12 (1.05-1.20)	$1.07x10^{-03}$	1.14 (1.09-1.18)	$3.38x10^{-11}$	0.66
rs16890979	4	9922167	SLC2A9	T	0.17	0.79 (0.74-0.83)	$3.19x10^{-16}$	0.73 (0.67-0.80)	$1.05x10^{-11}$	0.77 (0.74-0.81)	$5.45x10^{-26}$	0.19
rs16891234	4	9946163	SLC2A9	C	0.24	1.16 (1.11-1.22)	$6.06x10^{-10}$	1.13 (1.05-1.22)	$8.86x10^{-04}$	1.15 (1.11-1.20)	$2.72x10^{-12}$	0.57
rs1229984	4	100239319	ADH1B	T	0.03	1.51 (1.34-1.69)	$5.00x10^{-12}$	1.44 (1.21-1.72)	$4.77x10^{-05}$	1.49 (1.35-1.64)	$1.15x10^{-15}$	0.68
rs114791459	4	88591554	LOC105377323	A	0.02	1.42 (1.26-1.60)	$7.99x10^{-09}$	1.47 (1.22-1.77)	$5.47x10^{-05}$	1.43 (1.30-1.59)	$2.01x10^{-12}$	0.76
rs114580333	4	88790118	MEPE	A	0.02	1.44 (1.26-1.63)	$3.01x10^{-08}$	1.39 (1.15-1.69)	$9.18x10^{-04}$	1.42 (1.28-1.59)	$1.10x10^{-10}$	0.79
rs2078267	11	64334114	SLC22A11	C	0.47	1.16 (1.11-1.21)	$1.72x10^{-12}$	1.14 (1.07-1.22)	$6.27x10^{-05}$	1.15 (1.11-1.20)	$6.65x10^{-16}$	0.62

[1]Fixed-effects meta-analysis of the discovery GWAS and the replication analysis.
[2]Adjusted for age, sex, and 10 first PCs
A1, allele 1/effect allele; Bp, base pair position; Chr, chromosome; Freq, frequency; OR, Odds Ratios; SNP, single nucleotide polymorphism

Table 4.6 Association between GWAS lead SNPs and gout cases and asymptomatic hyperuricaemia controls with SU ≥7.0 mg/dl.

SNP	Chr	bp	Gene	A1	OR (95% CI)	P value[1]	aOR (95% CI)[2]	P value[1]
rs1260326	2	27730940	GCKR	T	1.08 (1.04-1.12)	$7.46x10^{-04}$	1.09 (1.05-1.13)	$1.80x10^{-04}$
rs2231142*	4	89052323	ABCG2	T	1.45 (1.39-1.52)	$1.05x10^{-52}$	1.48 (1.41-1.55)	$6.83x10^{-57}$
rs13120400*	4	89033527	ABCG2	C	0.87 (0.83-0.91)	$2.09x10^{-09}$	0.86 (0.83-0.90)	$1.11x10^{-10}$
rs7672194*	4	89126647	ABCG2	T	1.12 (1.08-1.17)	$5.93x10^{-09}$	1.13 (1.09-1.17)	$4.03x10^{-09}$
rs4693211*	4	89249061	PPM1K-DT	C	1.26 (1.18-1.35)	$8.85x10^{-11}$	1.29 (1.21-1.38)	$8.34x10^{-12}$
rs28793136*	4	89216768	PPM1K-DT	C	1.25 (1.18-1.34)	$6.30x10^{-11}$	1.26 (1.19-1.35)	$1.07x10^{-12}$
rs1545207	4	89239492	PPM1K-DT	A	1.11 (1.06-1.15)	$8.98x10^{-06}$	1.12 (1.07-1.16)	$9.29x10^{-07}$
rs16890979	4	9922167	SLC2A9	T	0.88 (0.84-0.93)	$3.73x10^{-05}$	0.85 (0.80-0.89)	$2.13x10^{-06}$
rs16891234	4	9946163	SLC2A9	C	1.10 (1.05-1.15)	$1.31x10^{-04}$	1.10 (1.05-1.15)	$1.50x10^{-04}$
rs1229984*	4	100239319	ADH1B	T	1.38 (1.24-1.53)	$1.43x10^{-08}$	1.38 (1.25-1.53)	$8.38x10^{-09}$
rs114791459*	4	88591554	LOC105377323	A	1.34 (1.21-1.50)	$9.03x10^{-07}$	1.35 (1.21-1.50)	$1.64x10^{-08}$
rs114580333	4	88790118	MEPE	A	1.33 (1.18-1.49)	$1.66x10^{-05}$	1.37 (1.22-1.53)	$9.39x10^{-07}$
rs2078267	11	64334114	SLC22A11	C	1.11 (1.07-1.15)	$9.49x10^{-07}$	1.11 (1.07-1.16)	$1.59x10^{-07}$

[1] p values were corrected for multiple testing using Bonferroni correction. [2] Adjusted for age and sex. *GWAS significant
A1, allele 1/effect allele; Bp, base pair position; Chr, chromosome; OR, Odds ratio; SNP, single nucleotide polymorphism; SU, serum urate.

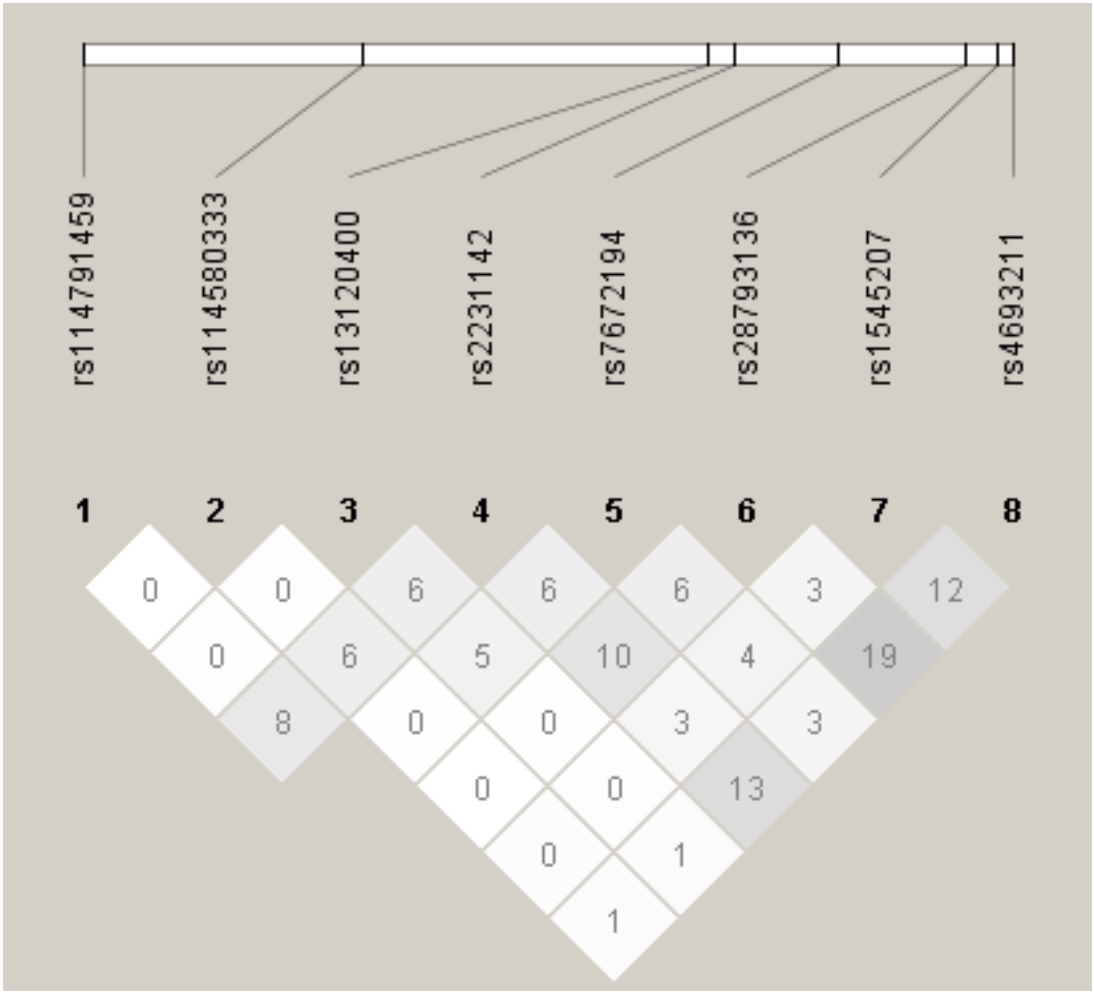


Figure 4.4 LD plot for regions in genes ABCG2, PPM1K-DT, MEPE and LOC105377323 on chromosome 4.

Each diamond represents r^2 values between two SNPs in European population. The plot was generated using HaploView 4.2.

4.3.3 Association of lead SNPs and serum urate

Linear regression was used to investigate whether the loci identified as independent associations with gout, had an effect on SU variation. This was conducted using the full cohort, and after adjusting for age and sex, and correcting for multiple testing, all lead SNPs associated with SU. rs2231142 in ABCG2 and rs16890979 in SLC2A9 showed the greatest effects: adjusted β=0.107 and $p=1.21x10^{-80}$, and adjusted β=-0.055 and $p=1.67x10^{-43}$, respectively (Table 4.7).

Table 4.7 Association between GWAS lead SNPs and serum urate levels.

SNP	Chr	bp	Gene	A1	β (SE)	P value[1]	Adjusted β (SE)[2]	P value[1]
rs1260326	2	27730940	GCKR	T	0.026 (0.004)	$8.68x10^{-10}$	0.028 (0.004)	$7.97x10^{-12}$
rs2231142	4	89052323	ABCG2	T	0.101 (0.006)	$3.86x10^{-71}$	0.107 (0.006)	$1.21x10^{-80}$
rs13120400	4	89033527	ABCG2	C	-0.037 (0.004)	$3.05x10^{-16}$	-0.039 (0.004)	$2.59x10^{-18}$
rs7672194	4	89126647	ABCG2	T	0.026 (0.004)	$7.55x10^{-10}$	0.027 (0.004)	$7.49x10^{-11}$
rs4693211	4	89249061	PPM1K-DT	C	0.050 (0.008)	$1.55x10^{-09}$	0.054 (0.008)	$3.27x10^{-11}$
rs28793136	4	89216768	PPM1K-DT	C	0.040 (0.007)	$8.11x10^{-07}$	0.042 (0.007)	$9.94x10^{-08}$
rs1545207	4	89239492	PPM1K-DT	A	0.020 (0.004)	$5.15x10^{-05}$	0.021 (0.004)	$9.08x10^{-06}$
rs16890979*	4	9922167	SLC2A9	T	-0.055 (0.005)	$7.59x10^{-26}$	-0.072 (0.005)	$1.67x10^{-43}$
rs16891234	4	9946163	SLC2A9	C	0.036 (0.005)	$1.11x10^{-13}$	0.035 (0.005)	$2.35x10^{-13}$
rs1229984	4	100239319	ADH1B	T	0.056 (0.013)	$1.40x10^{-04}$	0.055 (0.013)	$1.61x10^{-04}$
rs114791459	4	88591554	LOC105377323	A	0.056 (0.013)	$1.52x10^{-04}$	0.060 (0.013)	$3.71x10^{-05}$
rs114580333	4	88790118	MEPE	A	0.065 (0.014)	$2.59x10^{-05}$	0.070 (0.014)	$3.32x10^{-06}$
rs2078267	11	64334114	SLC22A11	C	0.029 (0.004)	$4.10x10^{-12}$	0.030 (0.004)	$1.66x10^{-13}$

[1] p values were corrected for multiple testing using Bonferroni correction. [2] Adjusted for age and sex. *Note that rs16890979 is in tight LD with the previously reported GWAS hit rs12498742 (r^2=0.79).
A1, allele 1/effect allele; Bp, base pair position; Chr, chromosome; SNP, single nucleotide polymorphism.

4.3.4 Polygenic risk score model for gout vs asymptomatic hyperuricaemia

A PRS for gout cases and asymptomatic hyperuricaemia controls was generated using PRSice, with the replication cohort as the test dataset and the GWAS summary statistics as the base dataset. Initially, 639,925 variants were included, and after applying the clumping parameters, a final number of 266,754 SNPs remained available for PRS calculation. PRSice calculated the scores at various significance thresholds, figure 4.5 illustrates the PRS model fit across the different p value thresholds. The best-fit p value threshold that gave the highest Nagelkerke's pseudo R^2 (0.016) was $4.0x10^{-6}$, and included 17 SNPs (Table 4.8).

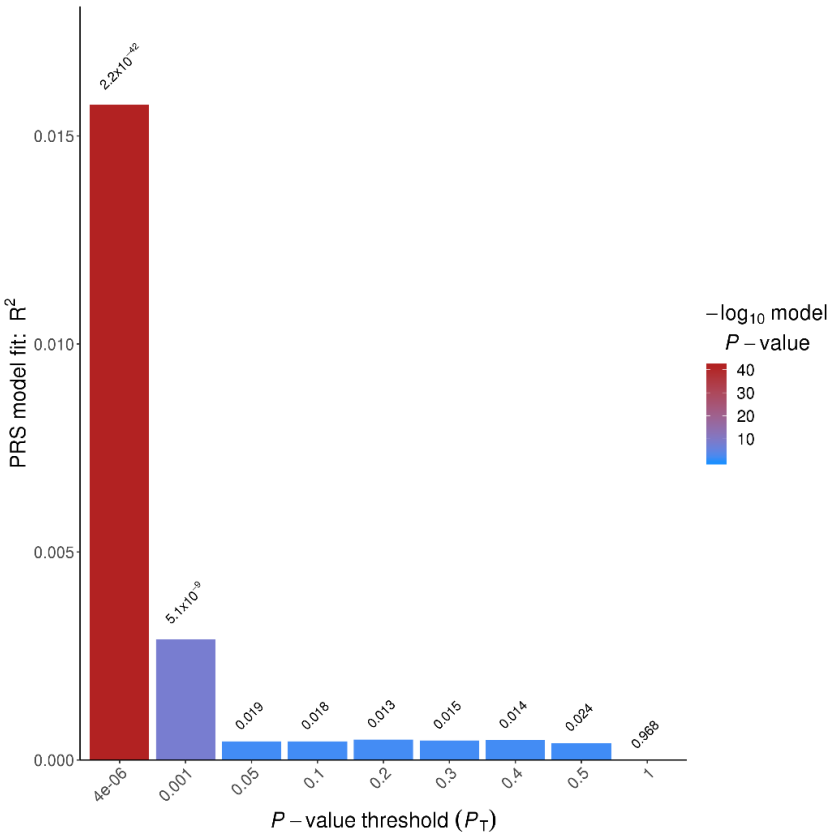


Figure 4.5 PRSice bar plot indicating the PRS model fit across different p value thresholds. The highest Nagelkerke's R^2 (0.016) was obtained with a p value=$4.0x10^{-6}$.

Table 4.8 SNPs under the best-fit p value threshold included into the PRS model.

SNP	Chr	bp	Gene	P value
rs1260326	2	27730940	GCKR	$3.54x10^{-08}$
rs2231142	4	89052323	ABCG2	$2.05x10^{-78}$
rs13120400	4	89033527	ABCG2	$1.56x10^{-16}$
rs7672194	4	89126647	ABCG2	$3.21x10^{-12}$
rs7658584	4	89096641	ABCG2	$3.27x10^{-06}$
rs28793136	4	89216768	PPM1K-DT	$8.19x10^{-17}$
rs1545207	4	89239492	PPM1K-DT	$6.82x10^{-09}$
rs16890979	4	9922167	SLC2A9	$3.19x10^{-16}$
rs16891234	4	9946163	SLC2A9	$6.06x10^{-10}$
rs1229984	4	100239319	ADH1B	$5.00x10^{-12}$
rs114791459	4	88591554	LOC105377323	$7.99x10^{-09}$
rs114580333	4	88790118	MEPE	$3.01x10^{-08}$
rs2078267	11	64334114	SLC22A11	$1.72x10^{-12}$
rs505802	11	64357072	SLC22A12	$1.43x10^{-09}$
rs2229357	12	57843711	INHBC	$2.18x10^{-06}$
rs7147313	14	23177202	LOC105370404	$2.61x10^{-06}$
rs55673000	17	38364764	MSL1	$2.41x10^{-06}$

Bp, base pair position; Chr, chromosome; SNP, single nucleotide polymorphism.

These SNPs included the GWAS hits located in or near GCKR, ABCG2, PPM1K-DT, SLC2A9, ADH1B, MEPE and LOC105377323, and four additional variants in SLC22A12, INHBC, MSL1 and LOC105370404 that did not reach genome wide significance in the discovery GWAS, but contributed to PRS best fit model.

Although there was a considerable overlap between the scores of cases and controls, the mean (±SD) PRS for gout cases was 0.018 (±0.017) compared to 0.013 (±0.016) for asymptomatic hyperuricaemia controls. The difference between groups was significant ($p < 0.0001$) (Figure 4.6).

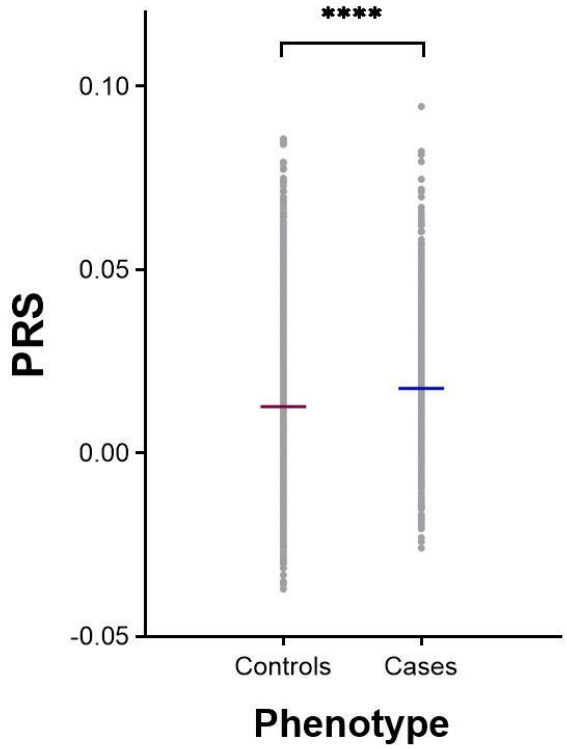


Figure 4.6 Distribution of polygenic risk scores (PRS) among gout cases and asymptomatic hyperuricaemia controls.

The horizontal lines represent the mean scores for both cases and controls (0.013 and 0.018, respectively), which showed a significant difference ($p<0.0001$).

The predictive ability of this PRS model was evaluated using the AUROC curve, and compared to the demographics model (age, sex and BMI) and combined model (age, sex, BMI and PRS). The AUC for the PRS model was 58.5%. The demographics model had a 66.7% predictive ability, and when the 17 variants when added to generate a combined model, the predictive ability increased to 69.2%. Figure 4.7 displays the ROC curves for each model.

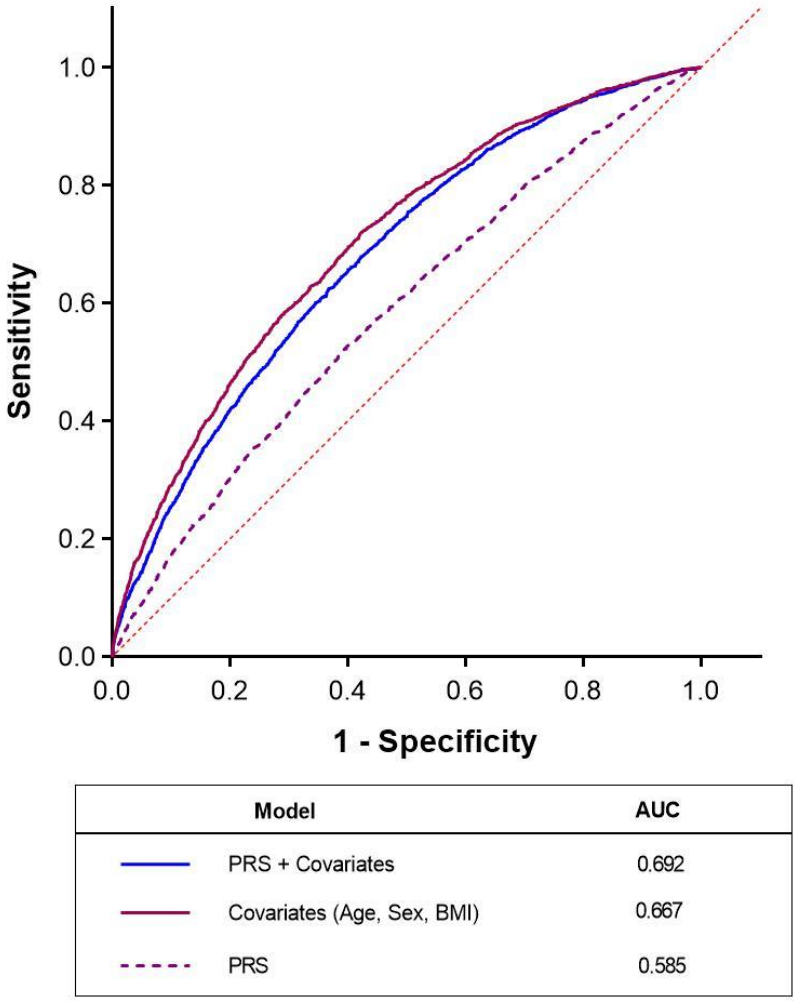


Figure 4.7 Area under the receiver operating characteristics (AUROC) curve for the PRS model, demographics model, and combined (demographics + PRS) model.

4.3.5 GWAS: Gout vs. normouricaemia

In order to explore if the lead SNPs associated with the transition from hyperuricaemia to gout, were associated with gout when compared to normouricaemia controls, two additional GWAS were conducted. These analyses involved gout cases vs. normouricaemia controls using a SU cut-off values of <6.0 mg/dL and <7.0 mg/dL. The first GWAS identified 52 independent SNPs, while the second identified 46 independent SNPs (Figure 4.8, Table 4.9). Figure 4.9 displays a graphical representation to compare the ORs for each locus in the GWAS of gout vs. asymptomatic hyperuricaemia and the GWAS of gout vs. normouricaemia (SU <6.0 mg/dL). Since several SNPs were present in the same gene, the variant with the smallest p-value was selected and plotted. To be consistent with the sensitivity analysis, a comparison of the OR of the thirteen lead SNPs upon exclusion of AH controls with SU 6.0-7.0 mg/dL, with the gout vs SU <7.0 mg/dL GWAS was also performed (Figure 4.9B). The same loci were responsible for transition from asymptomatic hyperuricaemia and normouricaemia to gout.

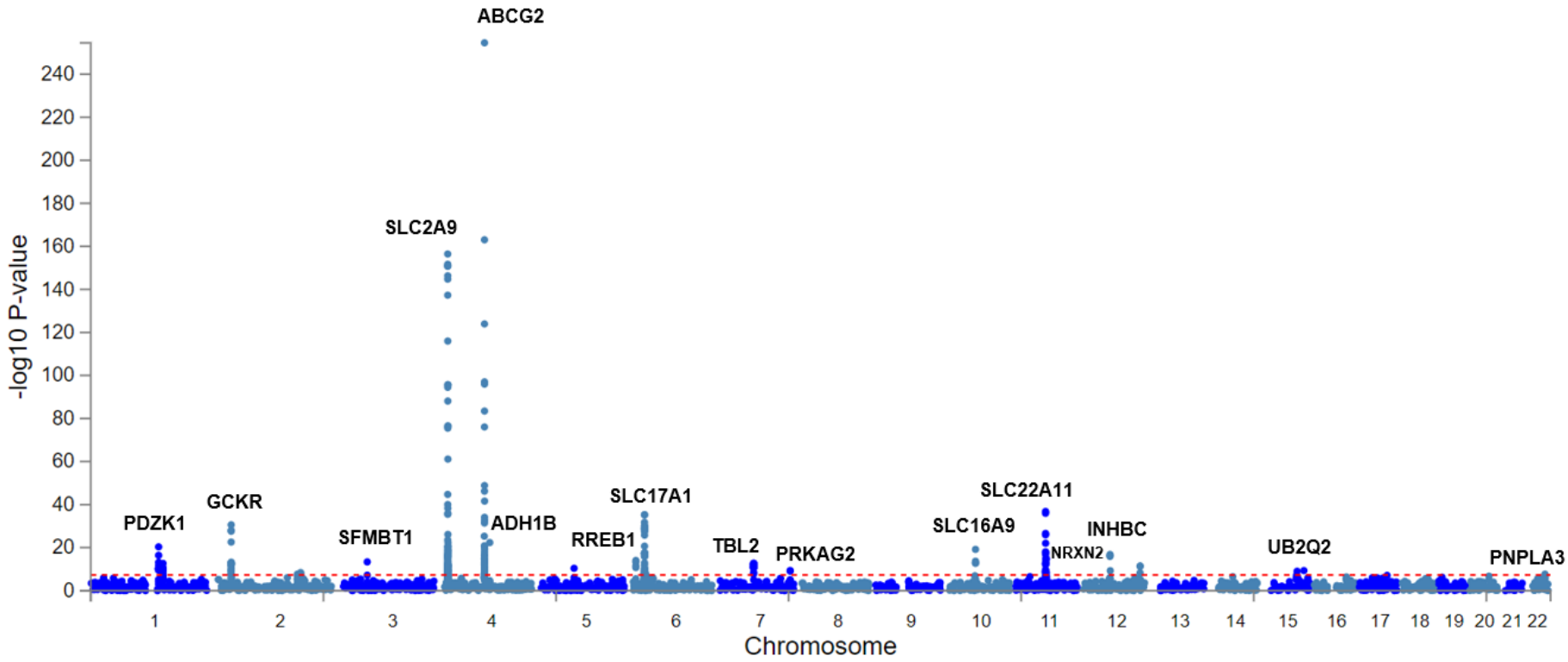


Figure 4.8 Manhattan plot of the discovery GWAS of gout vs. normouricaemia controls.

The y-axis shows –log10 P values ordered by chromosomal position on the x-axis. The horizontal dashed-line represents genome wide significance threshold (5.0×10^{-8}).

Table 4.9 Summary statistics of the GWAS for gout vs SU <6.0 mg/dL, and gout vs SU <7.0 mg/dL

						Gout vs SU <6.0 mg/dL		Gout vs SU <7.0 mg/dL	
SNP	**Chr**	**bp**	**Gene**	**A1**	**Freq.**	**aOR (95% CI)[1]**	**P value**	**aOR (95% CI)[1]**	**P value**
rs1967017	1	145723645	PDZK1	T	0.46	1.20 (1.16-1.25)	4.58E-21	1.17 (1.12-1.21)	1.27E-16
rs760077	1	155178782	MTX1	A	0.40	1.16 (1.11-1.20)	2.28E-13	1.13 (1.09-1.17)	5.40E-11
rs1260326	2	27730940	GCKR	T	0.39	1.26 (1.21-1.31)	2.55E-31	1.23 (1.18-1.28)	7.74E-28
rs2075252	2	170010985	LRP2	T	0.24	1.13 (1.08-1.18)	2.24E-08	NA[2]	NA
rs72929103	2	177047172	HAGLR	A	0.04	1.33 (1.21-1.47)	3.96E-09	NA	NA
rs2581778	3	53055311	SFMBT1	A	0.42	1.16 (1.12-1.20)	4.34E-14	1.14 (1.10-1.18)	5.04E-12
rs13140984	4	9637886	Intergenic	G	0.37	0.86 (0.83-0.90)	5.26E-13	0.89 (0.86-0.93)	5.43E-09
rs7698858	4	9895070	SLC2A9	C	0.25	0.81 (0.77-0.86)	6.03E-21	0.84 (0.81-0.88)	1.07E-14
rs75181558	4	9916133	SLC2A9	G	0.02	0.50 (0.43-0.59)	5.87E-17	0.55 (0.47-0.65)	2.65E-13
rs34213329	4	9917484	SLC2A9	G	0.02	0.56 (0.48-0.65)	7.25E-14	0.62 (0.53-0.72)	2.41E-10
rs13129697	4	9926967	SLC2A9	G	0.27	0.52 (0.50-0.55)	3.59E-157	0.57 (0.55-0.60)	2.54E-123
rs16891234	4	9946163	SLC2A9	C	0.22	1.35 (1.29-1.41)	9.87E-41	1.30 (1.25-1.36)	5.26E-35
rs75341455	4	10061147	WDR1	T	0.03	0.58 (0.51-0.66)	4.75E-17	0.63 (0.56-0.71)	4.86E-13
rs114695316	4	10291550	AC006499	A	0.03	0.57 (0.51-0.65)	5.61E-19	0.62 (0.55-0.70)	7.65E-15

Table 4.9 (Cont). Summary statistics of the GWAS for gout vs SU <6.0 mg/dL, and gout vs SU <7.0 mg/dL

						Gout vs SU <6.0 mg/dL		Gout vs SU <7.0 mg/dL	
SNP	**Chr**	**bp**	**Gene**	**A1**	**Freq.**	**aOR (95% CI)[1]**	**P value**	**aOR (95% CI)[1]**	**P value**
rs6839820	4	10296114	Intergenic	C	0.33	1.24 (1.19-1.29)	7.84E-27	1.20 (1.16-1.25)	5.82E-21
rs75816008	4	10335496	Intergenic	T	0.03	0.64 (0.57-0.72)	1.92E-13	0.70 (0.62-0.77)	2.25E-10
rs62286747	4	10370537	Intergenic	T	0.05	1.34 (1.23-1.45)	8.29E-13	1.29 (1.19-1.39)	4.10E-11
rs75487798	4	10445207	ZNF518B	C	0.04	1.30 (1.19-1.43)	4.39E-08	NA	NA
rs10016702	4	10447640	ZNF518B	G	0.10	1.19 (1.12-1.26)	2.53E-08	NA	NA
rs16869379	4	10514681	CLNK	C	0.31	0.81 (0.78-0.84)	5.17E-23	0.83 (0.80-0.86)	1.26E-19
rs114791459	4	88591554	LOC105377323	A	0.02	1.86 (1.66-2.09)	5.16E-26	1.72 (1.54-1.91)	6.68E-23
rs17013285	4	88767008	MEPE	A	0.17	1.22 (1.16-1.28)	1.54E-15	1.19 (1.14-1.25)	2.93E-13
rs114580333	4	88790118	MEPE	A	0.02	1.83 (1.62-2.07)	1.47E-21	1.70 (1.51-1.91)	3.18E-19
rs7676673	4	88869086	Intergenic	C	0.47	1.16 (1.11-1.20)	7.64E-14	1.14 (1.10-1.18)	1.30E-12
rs2728127	4	88895115	SPP1	G	0.29	0.83 (0.80-0.87)	3.20E-17	0.84 (0.80-0.87)	2.84E-17
rs2231142	4	89052323	ABCG2	T	0.12	2.50 (2.37-2.64)	2.61E-255	2.24 (2.13-2.35)	5.05E-231
rs2622604	4	89078924	ABCG2	T	0.26	0.71 (0.67-0.74)	1.42E-49	0.73 (0.69-0.76)	2.50E-45
rs7658584	4	89096641	ABCG2	A	0.15	0.81 (0.76-0.86)	2.57E-14	0.82 (0.78-0.87)	4.82E-13

Table 4.9 (Cont). Summary statistics of the GWAS for gout vs SU <6.0 mg/dL, and gout vs SU <7.0 mg/dL

						Gout vs SU <6.0 mg/dL		Gout vs SU <7.0 mg/dL	
SNP	**Chr**	**bp**	**Gene**	**A1**	**Freq.**	**aOR (95% CI)[1]**	**P value**	**aOR (95% CI)[1]**	**P value**
rs7672194	4	89126647	ABCG2	T	0.46	1.27 (1.22-1.32)	8.08E-35	1.25 (1.21-1.30)	3.43E-33
rs28793136	4	89216768	PPM1K-DT	C	0.07	1.51 (1.41-1.62)	2.24E-33	1.45 (1.36-1.55)	3.08E-30
rs1545207	4	89239492	PPM1K-DT	A	0.27	1.21 (1.16-1.27)	6.48E-20	1.20 (1.15-1.25)	6.96E-19
rs4693211	4	89249061	PPM1K-DT	C	0.06	1.64 (1.53-1.76)	2.45E-42	1.58 (1.48-1.69)	5.38E-41
rs1229984	4	100239319	ADH1B	T	0.02	1.75 (1.57-1.95)	4.63E-23	1.67 (1.50-1.85)	8.48E-22
rs520007	5	72445501	LOC105379030	C	0.43	1.14 (1.10-1.18)	4.17E-11	1.12 (1.08-1.16)	2.65E-09
rs11755724	6	7118990	RREB1	A	0.36	1.17 (1.12-1.21)	8.94E-15	1.14 (1.10-1.19)	1.75E-12
rs9461183	6	25510649	CARMIL1	G	0.42	1.12 (1.07-1.16)	3.09E-08	NA	NA
rs4712972	6	25772047	SLC17A1	A	0.15	1.21 (1.14-1.276)	1.34E-12	1.18 (1.12-1.24)	6.90E-11
rs1165196	6	25813150	SLC17A1	G	0.43	0.78 (0.75-0.81)	5.11E-36	0.81 (0.78-0.84)	2.15E-27
rs3800307	6	27185792	PRSS16	A	0.21	1.15 (1.10-1.20)	5.10E-09	1.13 (1.08-1.18)	3.50E-08
rs2286276	7	72987354	TBL2	T	0.28	0.85 (0.82-0.89)	1.94E-13	NA	NA
rs3812316	7	73020337	MLXIPL	G	0.13	NA	NA	0.81 (0.77-0.86)	2.15E-12
rs10224002	7	151415041	PRKAG2	G	0.29	1.14 (1.10-1.19)	4.45E-10	NA	NA

Table 4.9 (Cont). Summary statistics of the GWAS for gout vs SU <6.0 mg/dL, and gout vs SU <7.0 mg/dL

						Gout vs SU <6.0 mg/dL		Gout vs SU <7.0 mg/dL	
SNP	**Chr**	**bp**	**Gene**	**A1**	**Freq.**	**aOR (95% CI)[1]**	**P value**	**aOR (95% CI)[1]**	**P value**
rs1171619	10	61465838	SLC16A9	A	0.22	0.81 (0.76-0.84)	6.74E-20	0.83 (0.79-0.87)	4.85E-16
rs77085155	11	64154981	Intergenic	C	0.04	1.42 (1.29-1.56)	1.58E-13	1.38 (1.27-1.51)	6.36E-13
rs2078267	11	64334114	SLC22A11	C	0.45	1.28 (1.24-1.33)	1.68E-37	1.26 (1.21-1.30)	2.91E-34
rs55908299	11	64456148	NRXN2	A	0.07	0.79 (0.73-0.85)	6.95E-09	0.80 (0.74-0.86)	1.84E-08
rs471618	11	64465403	NRXN2	C	0.19	1.23 (1.18-1.29)	1.90E-18	1.22 (1.16-1.27)	1.46E-17
rs11227299	11	65549570	AP5B1	G	0.34	1.18 (1.13-1.22)	1.28E-15	NA	NA
rs948493	11	65552154	OVOL1	T	0.34	NA	NA	1.14 (1.10-1.19)	4.72E-12
rs2229357	12	57843711	INHBC	A	0.24	0.82 (0.78-0.86)	1.54E-17	0.83 (0.80-0.87)	2.20E-16
rs28548845	12	122594249	MLXIP	C	0.48	0.87 (0.84-0.91)	3.37E-12	NA	NA
rs7488857	12	122625212	MLXIP	A	0.48	NA	NA	0.88 (0.85-0.92)	4.16E-11
rs1394125	15	76158983	UBE2Q2	A	0.36	1.13 (1.09-1.18)	1.04E-09	1.11 (1.07-1.15)	3.53E-08
rs28508560	15	90642706	IDH2	T	0.23	1.15 (1.10-1.20)	4.76E-10	1.13 (1.08-1.18)	3.57E-08
rs738409	22	44324727	PNPLA3	G	0.22	0.87 (0.83-0.92)	2.17E-08	0.88 (0.84-0.92)	1.18E-08

A1, allele 1/effect allele; Bp, base pair position; Chr, chromosome; Freq, frequency; OR, Odds Ratios; SNP, single nucleotide polymorphism. [1] Adjusted for age, sex, and 10PCs.

[2]Not GWAS significant.

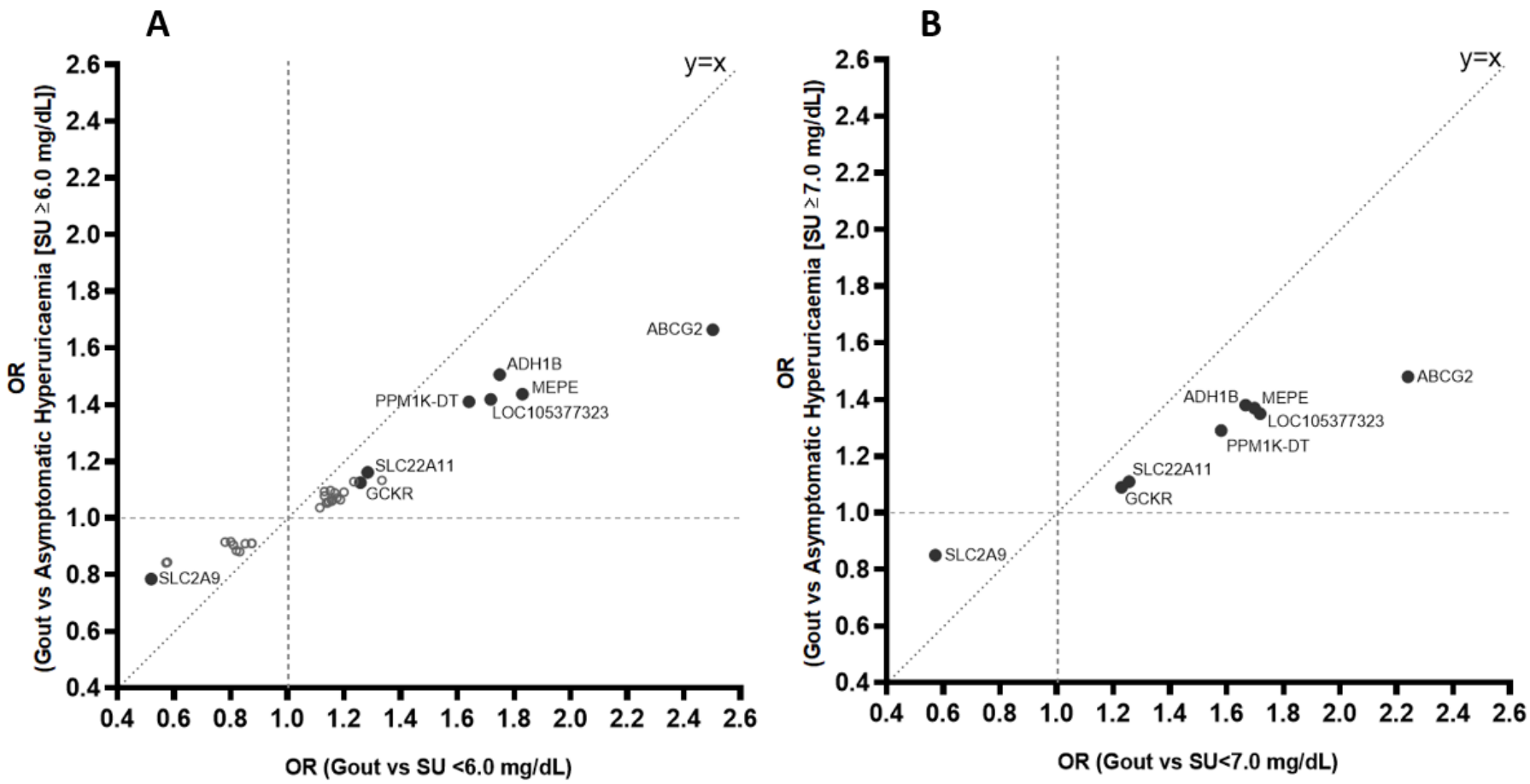


Figure 4.9 Scatter plots.

A) Comparison of the ORs of lead SNPs for both GWAS: gout vs asymptomatic hyperuricaemia (SU ≥6.0 mg/dL) and gout vs SU <6.0 mg/dL. Black dots represent ORs of the common risk loci of both GWAS, while grey circles represent ORs of additional lead SNPs of the gout vs SU<6.0 mg/dL that were not significant at GWAS level in the gout vs asymptomatic hyperuricaemia GWAS. B) Comparison of the ORs of the thirteen lead SNPs for gout vs asymptomatic hyperuricaemia (defined as SU ≥7.0 mg/dL), compared to the ORs in the GWAS for gout vs SU <7.0 mg/dL. Where several SNPs were present in the same gene, only that with the smallest p value was plotted in this graph.

4.4 Discussion

This study includes the largest GWAS to date and the first in Caucasians to examine the SNPs associated with the transition from asymptomatic hyperuricaemia to gout. Using UK Biobank data, it identified 13 independent SNPs from eight loci that reached genome wide significance and replicated successfully. These loci include urate transporters (i.e. ABCG2, SLC2A9 and SLC22A11), metabolic pathway genes (i.e. GCKR and ADH1B), and *MEPE* gene that regulates renal phosphate handling and skeletal mineralization (Rowe, 2004). Of the eight loci, ABCG2, SLC2A9, SLC22A11, PPM1K-DT, GCKR, and MEPE have been associated with gout or SU levels in previous GWAS within different populations (C.-J. Chen et al., 2018; Kottgen et al., 2013; Nakayama et al., 2017; Adrienne Tin et al., 2019). However, apart from ABCG2 and SLC2A9, the remaining genes have never been associated with the transition from asymptomatic hyperuricaemia to gout.

The only previous gout vs. asymptomatic hyperuricaemia GWAS was conducted in a Japanese population, and reported three novel loci: rs7927466 in CNTN5, rs9952962 in MIR302F, and a suggestive locus rs12980365 in ZNF724 that do not affect SU (Kawamura et al., 2019). Although rs7927466 is not included in the UK Biobank genotype platform, it is covered by its proxy SNP rs7942264 (r^2=1), but it did not show an association with gout; neither did rs12980365. MIR302F was not included in UK Biobank platform and further research on this gene is needed.

ADH1B was identified as a risk variant for gout vs. symptomatic hyperuricaemia. It has never previously been associated with gout in a GWAS – even when compared to general population. ADH1B encodes the β subunit of the class 1 alcohol dehydrogenase, and together with ALDH (aldehyde dehydrogenase)

enzymes, they mediate key steps in alcohol metabolism by the oxidation of ethanol into acetaldehyde by ADH1B, and acetaldehyde into acetate by ALDH2 (Polimanti & Gelernter, 2018). The SNP rs1229984 in ADH1B is functional variant that causes a change of an arginine to histidine. This increases ethanol clearance in liver, facilitates its conversion to highly reactive acetaldehyde (Edenberg, 2007), increases the NADH/NAD ratio that results in high lactic acid levels and increased urate reabsorption via URAT1 (Lieber, Jones, Losowsky, & Davidson, 1962). The risk allele of rs1229984 also promotes a "flush response" to alcohol and reduces the amount of alcohol consumed (Macgregor et al., 2009). Thus, the association between this polymorphism and gout may be due to increased production and reabsorption of urate from per unit alcohol consumed. This is consistent with the observation by Yokoyama *et al* in which rs1229984 associated with SU≥7 mg/dL (OR (95%CI) 2.04 (1.58–2.65)), while the daily alcohol intake was comparable across variants (Yokoyama et al., 2016). In agreement with these results, Sakiyama *et al* evaluated the effect of rs1229984 in ADH1B gene on gout (n=1,048 gout cases and 1,334 male controls). They reported an increased risk for gout with OR of 1.69 and 1.80 for His/Arg and His/His genotypes, respectively, that remained significant after correcting for alcohol consumption (M. Sakiyama et al., 2017). However, in their study patients with gout and rare variants of the SNP had greater alcohol consumption, suggesting an additional role for the latter. Further research is required to validate the association between rs1229984 and gout.

Urate transporters ABCG2, SLC2A9 and SLC22A11 play essential roles in pathogenesis of hyperuricaemia (Major et al., 2018; Woodward et al., 2009). SLC2A9 has the strongest effect on SU, accounting for 2-3% of variance, followed by ABCG2 that explains 1% of SU variation (Major et al., 2018). Although both loci have also been associated to gout, GWAS of gout cases vs

controls, have shown a greater effect of ABCG2 than SLC2A9 (Phipps-Green et al., 2016). In this study, rs2231142 in ABCG2 had a larger effect size on gout status compared to asymptomatic hyperuricaemia controls, than that of rs16890979 in SLC2A9 (which is in tight LD with the GWAS hit rs12498742) and also twice as much effect on SU than the latter. This supports the hypothesis that ABCG2 plays a causal role in the transition from hyperuricaemia to gout via its effect on SU. However, additional mechanisms may also operate. For instance, a smaller study using candidate gene approach, reported a nominal association for ABCG2 polymorphisms and gout vs. hyperuricaemia (Wrigley et al., 2020). It has been hypothesised that defects in ABCG2 cause a deficient autophagy, which is needed to form extracellular aggregates that resolve the inflammatory response to MSU crystals (Schorn et al., 2012; Wrigley et al., 2020). Additional research is needed to understand the exact mechanisms by which ABCG2 affects NLRP3-IL1β signalling.

The identified loci in PPM1K-DT and LOC105377323 are located in non-coding regions and their molecular mechanisms are unclear. Finally, MEPE may promote progression to gout via pro-mineralising osteopontin like function or via low phosphate levels that associates with incident hyperuricaemia (Cao et al., 2019).

The proximity between the variants located in or near LOC105377323, MEPE and PPM1K-DT and those located in ABCG2 in this study, raised the question of whether the first ones were genuine associations or were driven by the effect of rs2231142 in ABCG2. However, after clumping and conducting further pairwise LD comparisons, all the SNPs were retained as independent associations. Conditional analyses are often used as an approach to answer

these interrogations; nonetheless, they can also be too stringent and lead to ignoring loci with possible biological relevance.

ABCG2 gene has shown evidence of complex LD structure and additional SNPs within the same locus have been associated to gout. For instance, rs3114018 (r^2>0.2 with rs7672194, reported in our study), located in a different haplotype block from rs2231142, has been related to gout risk (K.-H. Yu et al., 2017). Additionally, a study by Zelenchuk *et al* reported that changes to the N-terminal region of MEPE contribute to hyperuricaemia and a reduction in fractional excretion of uric acid in a mouse model (Zelenchuk, Hedge, & Rowe, 2015). Therefore, the role of MEPE in purine metabolism merits additional research. Deep sequencing of the ABCG2 gene and up and downstream regions to capture *MEPE, PPM1K-DT* and *LOC105377323,* is required to determine underlying functional rare variants that could influence gout risk.

Regarding the PRS model, this is the first study that generates a PRS for predicting gout status in an asymptomatic hyperuricaemia population, and reported an AUC of 58.5% for genetic variants alone, that increased to 69.2% when demographic factors were added. In a previous study, Tin *et al* have generated a genetic risk score using variants associated with SU and examined their ability to predict gout cases in 334,800 UK Biobank participants not specifically selected for high SU levels. Their genetic risk score model had an AUC of 67% and 84% for the combined model (Adrienne Tin et al., 2019). The use of imputed data might improve the predictive ability, and is therefore suggested as future work.

The GWAS comparing gout cases with normouricaemic controls did not identify any inflammatory genes. A large number of lead SNPs were identified at genome-wide significance level. Most have been associated with gout or SU previously.(C.-J. Chen et al., 2018; Kottgen et al., 2013; H. Matsuo et al., 2013; Phipps-Green et al., 2016; Adrienne Tin et al., 2019) However, we identified three novel SNPs associated with gout compared to SU <6.0 mg/dL. Of these, rs11227299 (AP5B1) is associated with reduced eGFR (Pattaro et al., 2016). The variants in MTX1 and PRSS16 genes associated with gout compared to SU <6.0 mg/dL, also associate with Parkinson's Disease and Schizophrenia (Gan-Or, Bar-Shira, Gurevich, Giladi, & Orr-Urtreger, 2011; J. Shi et al., 2009). This is consistent with the negative associations between Parkinson's Disease and gout, and Schizophrenia and elevated SU (Alonso, Rodríguez, Logroscino, & Hernán, 2007; Q. He et al., 2020).

The strengths of this research include a large sample size, and the assessment of transition from asymptomatic hyperuricaemia or normouricaemia to gout in the same source population. However, there are several limitations. Firstly, gout definition was not based on ACR/EULAR classification criteria, but was ascertained via self-report of physician diagnosis, hospital diagnoses and ULT prescriptions. In addition, the classification of asymptomatic hyperuricaemia controls was based on a single SU measurement, which could have been affected by diet during the previous days. Additionally, the use of imputed was not possible due to computation time and storage capacity, and the use of non-imputed data limited the discovery power of the GWAS and PRS.

In conclusion, this study identified 13 GWAS significant risk loci, 12 of which have never previously been associated with the transition from asymptomatic

hyperuricaemia to gout at GWAS level. Larger GWAS are required to identify if variants in inflammatory pathways also contribute to this transition.

Chapter 5. Summary of the results and general discussion

5.1 Summary of aims and results

The general objective of this PhD project was to investigate genetic factors associated with the transition from asymptomatic hyperuricaemia to gout. This was conducted by two separate studies. The first one involved gene expression and cytokine measurements in participants recruited for different clinical studies at the Department of Academic Rheumatology. The aim was to evaluate differences among individuals with (1) normal SU levels (<360 µmol/L) without MSU crystal deposits; (2) high SU (≥360 µmol/L) without MSU crystal deposits; (3) high SU (≥360 µmol/L) with MSU crystal deposits; (4) intercritical gout; and (5) acute gout (Chapter 2). The second study involved the use of phenotype and genotype data from the UK Biobank. The UK Biobank is an extensive research resource that holds genetic and health-related information for >500,000 participants. Using genotype/phenotype data from the UK Biobank, we derived different phenotypes such as gout cases, non-gout controls and asymptomatic hyperuricaemia controls. First, we assemble a dataset of gout cases and controls to establish a robust QC pipeline to filter individuals and markers according to ethnicity, relatedness, heterozygosity, missingness and HWE. The resulting dataset was used to generate a PRS model to distinguish among gout cases and controls (Chapter 3). The same QC pipeline was used to construct an additional discovery dataset to perform a GWAS of gout cases vs. controls with asymptomatic hyperuricaemia. Association results were then validated using 30% of the original dataset, and used as the target dataset to construct a PRS model (Chapter 4). The key findings for each study are summarised in the following sections.

5.1.1 Study 1: Gene expression and cytokine analyses

Chapter 2 explored mRNA levels of 86 genes involved in the NLRP3 inflammasome and TLRs pathways in 108 individuals, classified according to SU levels, presence of MSU crystals or the stage of gout. Ten genes showed significant differences among the groups without gout, with *IL-1β* showing the greatest fold changes of 2.27 between SU ≥360 μmol/L and SU <360 μmol/L, and 3.79 between SU ≥360 μmol/L + MSU crystals and SU <360 μmol/L. In the comparison of intercritical gout vs. gout flare, no genes showed a differential expression after correcting for multiple testing. However, because the sample size was small (n=16), tentative differences were found for *CFLAR, NAIP, NFKBIA, NLRC4, NLRP6,* and *TLR2*, with NLRP6 showing the greatest fold change (0.42) in the gout flare group compared to the intercritical stage. We also examined the expression of the genes that showed significant differences in the previous analyses among the complete scenario of normouricaemia to acute gout. *CD40LG, PANX1* and *TNFSF14* showed a significant downregulation in the acute gout group compared to the SU ≥360μmol/L + MSU crystals group. These differences were unexpected, but further work to include a larger group of patients with gout is underway, which will allow us to verify these results.

The second stage of this study explored levels of 10 cytokines relevant to the immune response induced by MSU crystals in 185 participants. VEGF-α and hsCRP showed significant differences among the gout flare and the intercritical stage. And the levels of GRO-α, IL-1β, IL-6, IL-18, IP-10, MCP-1, TNF-α, and hsCRP were greater in participants with intercritical gout, compared to the non-gout participants. No significant differences were observed among the normouricaemia and the asymptomatic hyperuricaemia groups.

5.1.2 Study 2: GWAS and PRS models using data from the UK Biobank

The aims of this study were divided in two chapters. First, chapter 3 described the UK Biobank and our additional QC filters for genotype data. The resulting dataset was used to generate various cohorts: (1) one comprised by controls (matched by gender) and gout cases; (2) a cohort of gout cases and controls with asymptomatic hyperuricaemia; and (3) a cohort of gout cases and controls with normouricaemia.

The PRS for gout, described in chapter 3, was significantly associated with gout and provided a predictive ability of 62% that when added to the demographic variables, increased to a 75%. When exploring the risk variants that contributed to the scores, we found GWAS hits and additional loci that has been associated to gout; interestingly, one of those loci (*CLNK*) is involved in immune responses.

Chapter 4 describes the first GWAS in Caucasians to explore genetic variants in gout cases vs. asymptomatic hyperuricaemic controls. We reported 13 independent SNPs that replicated successfully in the validation cohort. These SNPs are located in or near *ABCG2, SLC2A9, SLC22A11, MEPE, GCKR*, *ADH1B*, and intronic regions (PPM1K-DT and LOC105377323). From these loci, ADH1B that plays an important role in alcohol metabolism, is a novel finding, as it has never been associated to gout or SU in previous GWAS. The PRS to distinguish between gout and asymptomatic hyperuricaemia had a predictive ability of 58.2% and 69.2% when combined with demographics data. Finally, the GWAS of gout vs. normouricaemia showed 52 and 46 independent SNPs for gout cases vs. normouricaemia controls with SU <6.0 mg/dL and SU <7.0 mg/dL, respectively. Three tentative novel associations were found in variants located in or near *AP5B1, MTX1* and *PRSS16*.

5.2 General discussion

In this research project we observed significant differences in NLRP3 and TLRs associated genes, but not in cytokines, in individuals with asymptomatic hyperuricaemia. Cytokine measurements showed elevated levels during the intercritical stage compared to normouricaemia and asymptomatic hyperuricaemia. Both observations could have an impact on the management of both asymptomatic hyperuricaemia and intercritical gout. The need to treat asymptomatic hyperuricemia has been widely debated. On one side, early treatments have been questioned due to their side effects like allopurinol hypersensitivity syndrome (Stamp & Dalbeth, 2017). However, more studies have associated asymptomatic hyperuricaemia with other cardiovascular comorbidities (Andrés et al., 2016; Miranda-Aquino et al., 2021), and have described that MSU crystals, even in the absence of obvious signs of acute inflammation, compromise joint function (Stewart et al., 2017). Additional research is needed to determine the true effects long-term of early interventions.

Conducting follow-up studies in larger sample sizes could assist in the discovery of additional mechanisms involved in gout, and in the identification of new biomarkers to differentiate each stage in the progression of the disease.

We could not associate the results from study 1 with the results from study 2, because the latter did not show any significant associations of genes involved in inflammatory mechanisms with the transition from asymptomatic hyperuricaemia to gout. This differs to the findings reported by Kawamura et al on their GWAS in Japanese population, were they reported two loci (CNTN5 and MIR302F) involved in immune responses that may aggravate asymptomatic hyperuricaemia into gout (Kawamura et al., 2019). There are several reasons

that might explain this. First, it could be that markers coverage for inflammatory-associated genes was low due to the use of directly genotyped data, rather than imputed genotypes. For instance, when exploring the UK Biobank array content for the 16 genes that showed significant differences among groups in the gene expression study (i.e. CLFAR, *IL-1β, NLRC4, NLRP6, TLR2, TXNIP,* etc.), only 10 were covered according to their location using the GRCh37 assembly, with less than 120 SNPs within those regions. Therefore, increasing SNPs coverage using imputed data could improve the discovery power. Using a pathway driven approach to generate PRS, can be used to evaluate what happens with the predictive ability to identify people with gout, when the model is restricted to inflammasome-associated loci. The use of a pathway driven approach has been reported recently for the prediction of late-onset Alzheimer's disease, were the PRS was restricted to synapse-encoding genes and the best model was based on eight SNPs that gave a predictive accuracy of 72% (Lawingco et al., 2021). Second, compared to GWAS for SU, those for gout are still limited in terms of sample sizes, and genetic variants exerting small effects are not likely to reach genome-wide significance. Additionally, GWAS assess associations of common SNPs (MAF >1%), which means that although if rare variants are impacting the disease risk, they would not be captured by GWAS (Visscher et al., 2017). Because of the limitations of GWAS, the use of sequencing technologies has increased as an alternative to fill the gaps in complex genomics research. In gout, recent studies that used whole-exome sequencing have revealed candidate rare variants located in *ABCG2, ADRB3, PRKG2* and *SLC16A9* in large pedigrees with gout (Huang et al., 2020; Xia et al., 2020).

Third, most of the risk loci identified by GWAS are located in non-coding regions, which hinders the understanding of their functional impact on the trait (Brodie, Azaria, & Ofran, 2016). Adding to this challenge, it is now known that SNPs

might also affect distant genes, which makes even more complicated the mapping of markers, because the nearest gene might not necessarily be the causal gene. An example of this is what occurred with *FTO* gene and obesity: GWAS reported that the strongest associations with increasing BMI were SNPs located in intronic regions of *FTO*; thus it was identified as an obesity-related gene (Scuteri et al., 2007). However, years later, additional studies revealed that those variants modified the expression of *IRX3*, which was the true link to BMI regulation (Smemo et al., 2014).

Finally, it is well known that genetic variation is also influenced by different mechanisms such as epigenetic regulation, chromatin remodelling and post-transcriptional modifications (van der Wijst, de Vries, Brugge, Westra, & Franke, 2018). Epigenetic modifications have been associated with inflammatory responses in patients with gout, specifically by the activation of micro-RNAs in the presence of MSU crystals (Nicola Dalbeth et al., 2015; H. M. Jin et al., 2014). Exploring these alternative mechanisms require different approaches, and integrating them is crucial to develop more robust fine-mapping algorithms to identify causal genes that would improve the understanding of complex diseases.

5.3 Conclusion

In this project, we explored gene expression and cytokine profiles among individuals with and without gout with varying SU levels and crystal deposition. We found significant differences that suggest the activation of inflammatory mechanisms in participants with asymptomatic hyperuricaemia, and systemic inflammation in intercritical gout. Additionally, our GWAS of gout vs. asymptomatic hyperuricaemia uncovered novel loci that has not been associated with gout before. With the UK Biobank-related work, we confirmed the preponderance of urate transporters and metabolic genes that affect SU levels, which supports the central role of hyperuricaemia in the pathogenesis of gout.

5.4 Future Research

One of the strengths of this research project was that participants with different stages of the disease were used to build our cohorts, and although the sample sizes were still limited, this type of well-characterised but larger datasets are needed to explore additional mechanisms involved in the transition from hyperuricaemia to gout and the severity of the disease. Both gene expression and genotype work conducted for this PhD project require further validation in independent datasets.

In terms of the discovery power, RT-PCR and GWAS studies present their own limitations, which have been described in previous chapters. Hence, sequencing approaches are suggested to identify additional variants that contribute to gout risk and to explore the associations of urate-related loci with gout. In this regard, the use of resources such as the UK Biobank will be of great value thanks to its large-scale genomic data that includes genotyping, whole-exome sequencing and whole-genome sequencing, plus the detailed phenotype information for >500,000 participants.

References

Abhishek, A., Courtney, P., Jenkins, W., Sandoval-Plata, G., Jones Adrian, C., Zhang, W., & Doherty, M. (2018). Monosodium urate crystal deposits are common in asymptomatic sons of people with gout - The Sons of gout study. *Arthritis & Rheumatology, 0*(ja). doi:10.1002/art.40572

Abhishek, A., Valdes, A. M., Jenkins, W., Zhang, W., & Doherty, M. (2017). Triggers of acute attacks of gout, does age of gout onset matter? A primary care based cross-sectional study. *PLoS One, 12*(10), e0186096. doi:10.1371/journal.pone.0186096

Akira, S., & Takeda, K. (2004). Toll-like receptor signalling. *Nature Reviews Immunology, 4*, 499. doi:10.1038/nri1391

Alberts, B. M., Bruce, C., Basnayake, K., Ghezzi, P., Davies, K. A., & Mullen, L. M. (2019). Secretion of IL-1β From Monocytes in Gout Is Redox Independent. *Frontiers in Immunology, 10*(70). doi:10.3389/fimmu.2019.00070

Allen, I. C., McElvania-TeKippe, E., Wilson, J. E., Lich, J. D., Arthur, J. C., Sullivan, J. T., . . . Ting, J. P. Y. (2013). Characterization of NLRP12 during the In Vivo Host Immune Response to Klebsiella pneumoniae and Mycobacterium tuberculosis. *PLoS One, 8*(4), e60842. doi:10.1371/journal.pone.0060842

Alonso, A., Rodríguez, L. A. G., Logroscino, G., & Hernán, M. A. (2007). Gout and risk of Parkinson disease. *Neurology, 69*(17), 1696. doi:10.1212/01.wnl.0000279518.10072.df

Andrés, M., Bernal, J.-A., Arenas, M.-D., & Pascual, E. (2019). Synovial fluid leukocyte count in asymptomatic hyperuricaemia with crystal deposition: a proof-of-concept study. *Rheumatology, 58*(6), 1104-1105. doi:10.1093/rheumatology/kez023

Andrés, M., Quintanilla, M.-A., Sivera, F., Sánchez-Payá, J., Pascual, E., Vela, P., & Ruiz-Nodar, J.-M. (2016). Silent Monosodium Urate Crystal Deposits Are Associated With Severe Coronary

Calcification in Asymptomatic Hyperuricemia: An Exploratory Study. *Arthritis & Rheumatology, 68*(6), 1531-1539. doi:10.1002/art.39581

Ansar, W., & Ghosh, S. (2013). C-reactive protein and the biology of disease. *Immunologic Research, 56*(1), 131-142. doi:10.1007/s12026-013-8384-0

Apte, R. S., Chen, D. S., & Ferrara, N. (2019). VEGF in Signaling and Disease: Beyond Discovery and Development. *Cell, 176*(6), 1248-1264. doi:10.1016/j.cell.2019.01.021

Balkarli, A., Tepeli, E., Balkarli, H., Kaya, A., & Cobankara, V. (2018). A variant allele of the Mediterranean-fever gene increases the severity of gout. *International Journal of Rheumatic Diseases, 21*(1), 338-346. doi:https://doi.org/10.1111/1756-185X.12872

Bardin, T. (2015). Canakinumab for the Patient With Difficult-to-Treat Gouty Arthritis: Review of the Clinical Evidence. *Joint Bone Spine, 82*, eS9-eS16. doi:https://doi.org/10.1016/S1297-319X(15)30003-8

Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2004). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics, 21*(2), 263-265. doi:10.1093/bioinformatics/bth457

Bayat, S., Baraf, H. S. B., & Rech, J. (2018). Update on imaging in gout: contrasting and comparing the role of dual-energy computed tomography to traditional diagnostic and monitoring techniques. (0392-856X (Print)).

Bonilla, F. A., & Oettgen, H. C. (2010). Adaptive immunity. *Journal of Allergy and Clinical Immunology, 125*(2, Supplement 2), S33-S40. doi:https://doi.org/10.1016/j.jaci.2009.09.017

Braga, T. T., Davanso, M. R., Mendes, D., de Souza, T. A., de Brito, A. F., Cruz, M. C., . . . Camara, N. O. S. (2021). Sensing soluble uric acid by Naip1-Nlrp3 platform. *Cell Death & Disease, 12*(2), 158. doi:10.1038/s41419-021-03445-w

Breuer, G. S., Bogot, N., & Nesher, G. (2016). Dual-energy computed tomography as a diagnostic tool for gout during intercritical

periods. *International Journal of Rheumatic Diseases, 19*(12), 1337-1341. doi:https://doi.org/10.1111/1756-185X.12938

Brodie, A., Azaria, J. R., & Ofran, Y. (2016). How far from the SNP may the causative genes be? *Nucleic Acids Research, 44*(13), 6046-6054. doi:10.1093/nar/gkw500

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., . . . Marchini, J. (2017). Genome-wide genetic data on ~500,000 UK Biobank participants. *bioRxiv*, 166298. doi:10.1101/166298

Cadzow, M., Merriman, T. R., & Dalbeth, N. (2017). Performance of gout definitions for genetic epidemiological studies: analysis of UK Biobank. *Arthritis Res Ther, 19*(1), 181. doi:10.1186/s13075-017-1390-1

Campion, E. W., Glynn, R. J., & DeLabry, L. O. (1987). Asymptomatic hyperuricemia. Risks and consequences in the Normative Aging Study. *Am J Med, 82*(3), 421-426.

Cao, J., Zhang, J., Li, Q., Jiang, C., Song, Y., Liu, C., . . . Qin, X. (2019). Serum Phosphate and the Risk of New-Onset Hyperuricemia in Hypertensive Patients. *Hypertension, 74*(1), 102-110. doi:10.1161/HYPERTENSIONAHA.119.12633

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience, 4*(1). doi:10.1186/s13742-015-0047-8

Chaudhury, S., Patel, T., Barber, I. S., Guetta-Baranes, T., Brookes, K. J., Chappell, S., . . . Morgan, K. (2018). Polygenic risk score in postmortem diagnosed sporadic early-onset Alzheimer's disease. *Neurobiology of Aging, 62*, 244.e241-244.e248. doi:https://doi.org/10.1016/j.neurobiolaging.2017.09.035

Chen-Xu, M., Yokose, C., Rai, S. K., Pillinger, M. H., & Choi, H. K. (2019). Contemporary Prevalence of Gout and Hyperuricemia in the United States and Decadal Trends: The National Health and Nutrition Examination Survey, 2007-2016. (2326-5205 (Electronic)).

Chen, C.-J., Tseng, C.-C., Yen, J.-H., Chang, J.-G., Chou, W.-C., Chu, H.-W., . . . Liao, W.-T. (2018). ABCG2 contributes to the development of gout and hyperuricemia in a genome-wide association study. *Scientific Reports, 8*(1), 3137. doi:10.1038/s41598-018-21425-7

Chen, C. J., Shi, Y., Hearn, A., Fitzgerald, K., Golenbock, D., Reed, G., . . . Rock, K. L. (2006). MyD88-dependent IL-1 receptor signaling is essential for gouty inflammation stimulated by monosodium urate crystals. *J Clin Invest, 116*(8), 2262-2271. doi:10.1172/JCI28075

Chen, Kaiwen W., Groß, Christina J., Sotomayor, Flor V., Stacey, Katryn J., Tschopp, J., Sweet, Matthew J., & Schroder, K. (2014). The Neutrophil NLRC4 Inflammasome Selectively Promotes IL-1β Maturation without Pyroptosis during Acute Salmonella Challenge. *Cell Reports, 8*(2), 570-582. doi:10.1016/j.celrep.2014.06.028

Chen, L., Wilson, J. E., Koenigsknecht, M. J., Chou, W.-C., Montgomery, S. A., Truax, A. D., . . . Ting, J. P. Y. (2017). NLRP12 attenuates colon inflammation by maintaining colonic microbial diversity and promoting protective commensal bacterial growth. *Nature Immunology, 18*(5), 541-551. doi:10.1038/ni.3690

Chhana, A., & Dalbeth, N. (2015). The gouty tophus: a review. *Curr Rheumatol Rep, 17*(3), 19. doi:10.1007/s11926-014-0492-x

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., Niu, J., Neogi, T., Chen, C. A., Chaisson, C., Hunter, D., & Zhang, Y. (2015). Nocturnal risk of gout attacks. *Arthritis & rheumatology (Hoboken, N.J.), 67*(2), 555-562. doi:10.1002/art.38917

Choi, H. K., Soriano, L. C., Zhang, Y., & Rodríguez, L. A. G. (2012). Antihypertensive drugs and risk of incident gout among patients with hypertension: population based case-control study. *BMJ, 344*.

Choi, S. W., Mak, T. S.-H., & O'Reilly, P. F. (2020). Tutorial: a guide to performing polygenic risk score analyses. *Nature Protocols, 15*(9), 2759-2772. doi:10.1038/s41596-020-0353-1

Choi, S. W., & O'Reilly, P. F. (2019). PRSice-2: Polygenic Risk Score software for biobank-scale data. *GigaScience, 8*(7). doi:10.1093/gigascience/giz082

Chowalloor, P., Raymond, W. D., Cheah, P., & Keen, H. (2020). The burden of subclinical intra-articular inflammation in gout. *International Journal of Rheumatic Diseases, 23*(5), 661-668. doi:https://doi.org/10.1111/1756-185X.13811

Chung, Y., Chang, S. H., Martinez, G. J., Yang, X. O., Nurieva, R., Kang, H. S., . . . Dong, C. (2009). Critical Regulation of Early Th17 Cell Differentiation by Interleukin-1 Signaling. *Immunity, 30*(4), 576-587. doi:10.1016/j.immuni.2009.02.007

Cicchese, J. M., Evans, S., Hult, C., Joslyn, L. R., Wessler, T., Millar, J. A., . . . Kirschner, D. E. (2018). Dynamic balance of pro- and anti-inflammatory signals controls disease and limits pathology. *Immunological Reviews, 285*(1), 147-167. doi:https://doi.org/10.1111/imr.12671

Clive, D. M. (2000). Renal Transplant-Associated Hyperuricemia and Gout. *Journal of the American Society of Nephrology, 11*(5), 974. doi:10.1681/ASN.V115974

Conforti-Andreoni, C., Spreafico, R., Qian, H. L., Riteau, N., Ryffel, B., Ricciardi-Castagnoli, P., & Mortellaro, A. (2011). Uric Acid-Driven Th17 Differentiation Requires Inflammasome-Derived IL-1 and IL-18. *The Journal of Immunology, 187*(11), 5842.

Dalbeth, N., & Doyle, A. J. (2012). Imaging of gout – An overview. *Best Practice & Research Clinical Rheumatology, 26*(6), 823-838. doi:https://doi.org/10.1016/j.berh.2012.09.003

Dalbeth, N., & Haskard, D. O. (2005). Mechanisms of inflammation in gout. *Rheumatology (Oxford), 44*(9), 1090-1096. doi:10.1093/rheumatology/keh640

Dalbeth, N., House, M. E., Aati, O., Tan, P., Franklin, C., Horne, A., . . . McQueen, F. M. (2015). Urate crystal deposition in asymptomatic hyperuricaemia and symptomatic gout: a dual energy CT study. *Ann Rheum Dis, 74*(5), 908-911. doi:10.1136/annrheumdis-2014-206397

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., 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 Diseases, 77*(7), 1048. doi:10.1136/annrheumdis-2017-212288

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. *Ann Rheum Dis, 77*(7), 1048-1052. doi:10.1136/annrheumdis-2017-212288

Dalbeth, N., Pool, B., Shaw, O. M., Harper, J. L., Tan, P., Franklin, C., . . . Naot, D. (2015). Role of miR-146a in regulation of the acute inflammatory response to monosodium urate crystals. *Annals of the Rheumatic Diseases, 74*(4), 786. doi:10.1136/annrheumdis-2014-205409

Dalbeth, N., & Stamp, L. (2014). Hyperuricaemia and gout: time for a new staging system? *Annals of the Rheumatic Diseases, 73*(9), 1598. doi:10.1136/annrheumdis-2014-205304

Dang, W., Xu, D., Xie, W., & Zhou, J. (2018). Study on the expressions of NLRP3 gene transcript variants in peripheral blood monocytes of primary gout patients. *Clinical Rheumatology, 37*(9), 2547-2555. doi:10.1007/s10067-018-4149-4

Das, S., Ghosh, A., Ghosh, P., Lahiri, D., Sinhamahapatra, P., & Basu, K. (2017). Sensitivity and specificity of ultrasonographic features of gout in intercritical and chronic phase. *International Journal of*

Rheumatic Diseases, 20(7), 887-893. doi:https://doi.org/10.1111/1756-185X.12928

de Ferranti, S., & Rifai, N. (2002). C-reactive protein and cardiovascular disease: a review of risk prediction and interventions. *Clinica Chimica Acta, 317*(1), 1-15. doi:https://doi.org/10.1016/S0009-8981(01)00797-5

De Miguel, E., Puig, J. G., Castillo, C., Peiteado, D., Torres, R. J., & Martin-Mola, E. (2012). Diagnosis of gout in patients with asymptomatic hyperuricaemia: a pilot ultrasound study. *Ann Rheum Dis, 71*(1), 157-158. doi:10.1136/ard.2011.154997

Dehghan, A., Köttgen, A., Yang, Q., Hwang, S.-J., Kao, W. H. L., Rivadeneira, F., . . . Fox, C. S. (2008). Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. *The Lancet, 372*(9654), 1953-1961. doi:10.1016/s0140-6736(08)61343-4

Derosa, G., Maffioli, P., & Sahebkar, A. (2015). Plasma uric acid concentrations are reduced by fenofibrate: A systematic review and meta-analysis of randomized placebo-controlled trials. *Pharmacological Research, 102*, 63-70. doi:https://doi.org/10.1016/j.phrs.2015.09.012

Dessein, P. H., Shipton, E. A., Stanwix, A. E., Joffe, B. I., & Ramokgadi, J. (2000). Beneficial effects of weight loss associated with moderate calorie/carbohydrate restriction, and increased proportional intake of protein and unsaturated fat on serum urate and lipoprotein levels in gout: a pilot study. *Annals of the Rheumatic Diseases, 59*(7), 539.

Diaz-Torne, C., Ortiz, M. A., Garcia-Guillen, A., Jeria-Navarro, S., Sainz, L., Fernandez-Sanchez, S., . . . Vidal, S. (2021). The inflammatory role of silent urate crystal deposition in intercritical gout. *Rheumatology*. doi:10.1093/rheumatology/keab335

Dinesh, P., & Rasool, M. (2017). Berberine, an isoquinoline alkaloid suppresses TXNIP mediated NLRP3 inflammasome activation in MSU crystal stimulated RAW 264.7 macrophages through the

upregulation of Nrf2 transcription factor and alleviates MSU crystal induced inflammation in rats. *International Immunopharmacology, 44*, 26-37. doi:https://doi.org/10.1016/j.intimp.2016.12.031

Dong, Z., Zhou, J., Xu, X., Jiang, S., Li, Y., Zhao, D., . . . Wang, J. (2018). Genetic variants in two pathways influence serum urate levels and gout risk: a systematic pathway analysis. *Scientific Reports, 8*(1), 3848. doi:10.1038/s41598-018-21858-0

Drivelegka, P., Sigurdardottir, V., Svärd, A., Jacobsson, L. T. H., & Dehlin, M. (2018). Comorbidity in gout at the time of first diagnosis: sex differences that may have implications for dosing of urate lowering therapy. *Arthritis Research & Therapy, 20*(1), 108. doi:10.1186/s13075-018-1596-x

Duncan, L., Shen, H., Gelaye, B., Meijsen, J., Ressler, K., Feldman, M., . . . Domingue, B. (2019). Analysis of polygenic risk score usage and performance in diverse human populations. *Nature Communications, 10*(1), 3328. doi:10.1038/s41467-019-11112-0

Edenberg, H. J. (2007). The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol research & health : the journal of the National Institute on Alcohol Abuse and Alcoholism, 30*(1), 5-13.

Edfors, F., Danielsson, F., Hallström, B. M., Käll, L., Lundberg, E., Pontén, F., . . . Uhlén, M. (2016). Gene-specific correlation of RNA and protein levels in human cells and tissues. *Molecular systems biology, 12*(10), 883-883. doi:10.15252/msb.20167144

Enomoto, A., Kimura, H., Chairoungdua, A., Shigeta, Y., Jutabha, P., Ho Cha, S., . . . Endou, H. (2002). Molecular identification of a renal urate–anion exchanger that regulates blood urate levels. *Nature, 417*, 447. doi:10.1038/nature742

https://www.nature.com/articles/nature742#supplementary-information

Estevez-Garcia, I. O., Gallegos-Nava, S., Vera-Pérez, E., Silveira, L. H., Ventura-Ríos, L., Vancini, G., . . . Amezcua-Guerra, L. M. (2018). Levels of Cytokines and MicroRNAs in Individuals With Asymptomatic Hyperuricemia and Ultrasonographic Findings of

Gout: A Bench-to-Bedside Approach. *Arthritis Care & Research, 70*(12), 1814-1821. doi:https://doi.org/10.1002/acr.23549

Euesden, J., Lewis, C. M., & O'Reilly, P. F. (2014). PRSice: Polygenic Risk Score software. *Bioinformatics, 31*(9), 1466-1468. doi:10.1093/bioinformatics/btu848

Filippucci, E., Reginato, A. M., & Thiele, R. G. (2020). Imaging of crystalline arthropathy in 2020. *Best Practice & Research Clinical Rheumatology, 34*(6), 101595. doi:https://doi.org/10.1016/j.berh.2020.101595

FitzGerald, J. D., Dalbeth, N., Mikuls, T., Brignardello-Petersen, R., Guyatt, G., Abeles, A. M., . . . Neogi, T. (2020). 2020 American College of Rheumatology Guideline for the Management of Gout. *Arthritis Care & Research, 72*(6), 744-760. doi:https://doi.org/10.1002/acr.24180

Franchi, L., Kamada, N., Nakamura, Y., Burberry, A., Kuffa, P., Suzuki, S., . . . Núñez, G. (2012). NLRC4-driven production of IL-1β discriminates between pathogenic and commensal bacteria and promotes host intestinal defense. *Nature Immunology, 13*(5), 449-456. doi:10.1038/ni.2263

Gan-Or, Z., Bar-Shira, A., Gurevich, T., Giladi, N., & Orr-Urtreger, A. (2011). Homozygosity for the MTX1 c.184T>A (p.S63T) alteration modifies the age of onset in GBA-associated Parkinson's disease. *neurogenetics, 12*(4), 325-332. doi:10.1007/s10048-011-0293-6

Garrod, A. B. (1848). Observations on certain pathological conditions of the blood and urine, in gout, rheumatism, and Bright's disease. (0959-5287 (Print)).

Genes, N., & Chisolm-Straker, M. (2012). Monoarticular arthritis update: Current evidence for diagnosis and treatment in the emergency department. *Emergency Medicine Practice, 14*(5).

Gentili, A. (2003). Advanced Imaging of Gout. *Semin Musculoskelet Radiol, 7*(03), 165-174.

Ghaemi-Oskouie, F., & Shi, Y. (2011). The Role of Uric Acid as an Endogenous Danger Signal in Immunity and Inflammation.

Current Rheumatology Reports, 13(2), 160-166. doi:10.1007/s11926-011-0162-1

Giles, D. A., Zahner, S., Krause, P., Van Der Gracht, E., Riffelmacher, T., Morris, V., . . . Kronenberg, M. (2018). The Tumor Necrosis Factor Superfamily Members TNFSF14 (LIGHT), Lymphotoxin β and Lymphotoxin β Receptor Interact to Regulate Intestinal Inflammation. *Frontiers in Immunology, 9*(2585). doi:10.3389/fimmu.2018.02585

Gordon, C., Swan, A., & Dieppe, P. (1989). Detection of crystals in synovial fluids by light microscopy: sensitivity and reliability. *Annals of the Rheumatic Diseases, 48*, 737-742.

Guo, H., Callaway, J. B., & Ting, J. P. (2015). Inflammasomes: mechanism of action, role in disease, and therapeutics. *Nat Med, 21*(7), 677-687. doi:10.1038/nm.3893

Gurung, P., Lukens, J. R., & Kanneganti, T.-D. (2015). Mitochondria: diversity in the regulation of the NLRP3 inflammasome. *Trends in Molecular Medicine, 21*(3), 193-201. doi:https://doi.org/10.1016/j.molmed.2014.11.008

Gutierrez, M., Schmidt, W. A., Thiele, R. G., Keen, H. I., Kaeley, G. S., Naredo, E., . . . group, O. U. G. T. F. (2015). International Consensus for ultrasound lesions in gout: results of Delphi process and web-reliability exercise. *Rheumatology (Oxford), 54*(10), 1797-1805. doi:10.1093/rheumatology/kev112

Hanson, L. Å., & Wigzell, H. (2014). *Immunology*. London, UNITED KINGDOM: Elsevier Science & Technology.

Harrold, L. R., Mazor, K. M., Negron, A., Ogarek, J., Firneno, C., & Yood, R. A. (2013). Primary care providers' knowledge, beliefs and treatment practices for gout: results of a physician questionnaire. *Rheumatology, 52*, 1623-1629. doi:doi:10.1093/rheumatology/ket158

He, Q., You, Y., Yu, L., Yao, L., Lu, H., Zhou, X., . . . Zhao, X. (2020). Uric acid levels in subjects with schizophrenia: A systematic

review and meta-analysis. *Psychiatry Research, 292*, 113305. doi:https://doi.org/10.1016/j.psychres.2020.113305

He, Y., Zeng, M. Y., Yang, D., Motro, B., & Nunez, G. (2016). NEK7 is an essential mediator of NLRP3 activation downstream of potassium efflux. *Nature, 530*(7590), 354-357. doi:10.1038/nature16959

Holzinger, D., Nippe, N., Vogl, T., Marketon, K., Mysore, V., Weinhage, T., . . . Roth, J. (2014). Myeloid-Related Proteins 8 and 14 Contribute to Monosodium Urate Monohydrate Crystal–Induced Inflammation in Gout. *Arthritis & Rheumatology, 66*(5), 1327-1339. doi:https://doi.org/10.1002/art.38369

Hu, Y., Li, L., Ehm, Margaret G., Bing, N., Song, K., Nelson, Matthew R., . . . Kang, Hyun M. (2013). The Benefits of Using Genetic Information to Design Prevention Trials. *The American Journal of Human Genetics, 92*(4), 547-557. doi:10.1016/j.ajhg.2013.03.003

Huang, X.-F., Sun, L., Zhang, C., Zhou, Z., Chen, H., Zhang, L., . . . Xia, X. (2020). Whole-Exome Sequencing Reveals a Rare Missense Variant in SLC16A9 in a Pedigree with Early-Onset Gout. *BioMed Research International, 2020*, 4321419. doi:10.1155/2020/4321419

Hui, M., Carr, A., Cameron, S., Davenport, G., Doherty, M., Forrester, H., . . . Guidelines Working, G. (2017). The British Society for Rheumatology Guideline for the Management of Gout. *Rheumatology, 56*(7), e1-e20. doi:10.1093/rheumatology/kex156

Hwang, J., Suh, H.-W., Jeon, Y. H., Hwang, E., Nguyen, L. T., Yeom, J., . . . Kim, M. H. (2014). The structural basis for the negative regulation of thioredoxin by thioredoxin-interacting protein. *Nature Communications, 5*, 2958. doi:10.1038/ncomms3958

https://www.nature.com/articles/ncomms3958#supplementary-information

Ichida, K., Matsuo, H., Takada, T., Nakayama, A., Murakami, K., Shimizu, T., . . . Suzuki, H. (2012). Decreased extra-renal urate excretion is a common cause of hyperuricemia. *Nat Commun, 3*, 764. doi:10.1038/ncomms1756

J Roessler, B., M Nosal, J., Smith, P., Heidler, S., D Palella, T., L Switzer, R., & Becker, M. (1994). *Human X-linked phosphoribosylpyrophosphate synthetase superactivity is associated with distinct point mutations in the PRPS1 gene* (Vol. 268).

Jamilloux, Y., Lefeuvre, L., Magnotti, F., Martin, A., Benezech, S., Allatif, O., . . . Henry, T. (2018). Familial Mediterranean fever mutations are hypermorphic mutations that specifically decrease the activation threshold of the Pyrin inflammasome. *Rheumatology, 57*(1), 100-111. doi:10.1093/rheumatology/kex373

Jansen, T. L., & Rasker, J. J. (2011). Therapeutic consequences of crystals in the synovial fluid: a review for clinicians. *Clin Exp Rheumatol, 29*(6), 1032-1039.

Jeck, W. R., Siebold, A. P., & Sharpless, N. E. (2012). Review: A Meta-Analysis of GWAS Studies and Age-Associated Diseases. *Aging cell, 11*(5), 727-731. doi:10.1111/j.1474-9726.2012.00871.x

Ji, A., Shaukat, A., Takei, R., Bixley, M., Cadzow, M., Topless, R. K., . . . Merriman, T. R. (2021). Aotearoa New Zealand Māori and Pacific population-specific gout risk variants: CLNK is a separate risk gene at the SLC2A9 locus. *The Journal of Rheumatology*, jrheum.201684. doi:10.3899/jrheum.201684

Jin, H. M., Kim, T.-J., Choi, J.-H., Kim, M.-J., Cho, Y.-N., Nam, K.-I., . . . Park, Y.-W. (2014). MicroRNA-155 as a proinflammatory regulator via SHIP-1 down-regulation in acute gouty arthritis. *Arthritis Research & Therapy, 16*(2), R88. doi:10.1186/ar4531

Jin, T.-b., Ren, Y., Shi, X., Jiri, M., He, N., Feng, T., . . . Kang, L. (2015). Genetic variations in the CLNK gene and ZNF518B gene are associated with gout in case–control sample sets. *Rheumatology International, 35*(7), 1141-1147. doi:10.1007/s00296-015-3215-3

Jinnah, H. A., Sabina, R. l., & Van Den Berghe, G. (2013). Chapter 187 - Metabolic disorders of purine metabolism affecting the nervous system. In O. Dulac, M. Lassonde, & H. B. Sarnat (Eds.),

Handbook of Clinical Neurology (Vol. 113, pp. 1827-1836): Elsevier.

Jo, E. K., Kim, J. K., Shin, D. M., & Sasakawa, C. (2016). Molecular mechanisms regulating NLRP3 inflammasome activation. *Cell Mol Immunol, 13*(2), 148-159. doi:10.1038/cmi.2015.95

Kaneko, K., & Maru, M. (2000). Determination of Urate Crystal Formation Using Flow Cytometry and Microarea X-Ray Diffractometry. *Analytical Biochemistry, 281*(1), 9-14. doi:https://doi.org/10.1006/abio.2000.4543

Kanevets, U., Sharma, K., Dresser, K., & Shi, Y. (2009). A Role of IgM Antibodies in Monosodium Urate Crystal Formation and Associated Adjuvanticity. *The Journal of Immunology, 182*(4), 1912.

Kawamura, Y., Nakaoka, H., Nakayama, A., Okada, Y., Yamamoto, K., Higashino, T., . . . Matsuo, H. (2019). Genome-wide association study revealed novel loci which aggravate asymptomatic hyperuricaemia into gout. *Annals of the Rheumatic Diseases, 78*(10), 1430. doi:10.1136/annrheumdis-2019-215521

Khanna, P. P., Nuki, G., Bardin, T., Tausche, A.-K., Forsythe, A., Goren, A., . . . Khanna, D. (2012). 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*(1), 117. doi:10.1186/1477-7525-10-117

Kienhorst, L. B. E., van Lochem, E., Kievit, W., Dalbeth, N., Merriman, M. E., Phipps-Green, A., . . . Radstake, T. R. D. J. (2015). Gout Is a Chronic Inflammatory Disease in Which High Levels of Interleukin-8 (CXCL8), Myeloid-Related Protein 8/Myeloid-Related Protein 14 Complex, and an Altered Proteome Are Associated With Diabetes Mellitus and Cardiovascular Disease. *Arthritis & Rheumatology, 67*(12), 3303-3313. doi:https://doi.org/10.1002/art.39318

King, D. E., Mainous, A. G., Buchanan, T. A., & Pearson, W. S. (2003). C-Reactive Protein and Glycemic Control in Adults With Diabetes. *Diabetes Care, 26*(5), 1535. doi:10.2337/diacare.26.5.1535

Kingsbury, S. R., Conaghan, P. G., & McDermott, M. F. (2011). The role of the NLRP3 inflammasome in gout. *J Inflamm Res, 4*, 39-49. doi:10.2147/JIR.S11330

Kolz, M., Johnson, T., Sanna, S., Teumer, A., Vitart, V., Perola, M., . . . Gieger, C. (2009). Meta-Analysis of 28,141 Individuals Identifies Common Variants within Five New Loci That Influence Uric Acid Concentrations. *PLOS Genetics, 5*(6), e1000504. doi:10.1371/journal.pgen.1000504

Kotake, S., Udagawa, N., Takahashi, N., Matsuzaki, K., Itoh, K., Ishiyama, S., . . . Suda, T. (1999). IL-17 in synovial fluids from patients with rheumatoid arthritis is a potent stimulator of osteoclastogenesis. *The Journal of Clinical Investigation, 103*(9), 1345-1352. doi:10.1172/JCI5703

Kottgen, A., Albrecht, E., Teumer, A., Vitart, V., Krumsiek, J., Hundertmark, C., . . . Gieger, C. (2013). Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nat Genet, 45*(2), 145-154. doi:10.1038/ng.2500

Kuo, C. F., Grainge, M. J., Mallen, C., Zhang, W., & Doherty, M. (2015). Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. *Ann Rheum Dis, 74*(4), 661-667. doi:10.1136/annrheumdis-2013-204463

Langford, H. G., Blaufox, M. D., Borhani, N. O., Curb, J. D., Molteni, A., Schneider, K. A., & Pressel, S. (1987). Is thiazide-produced uric acid elevation harmful? Analysis of data from the Hypertension Detection and Follow-up Program. *Archives of Internal Medicine, 147*(4), 645-649.

Larussa, T., Leone, I., Suraci, E., Imeneo, M., & Luzza, F. (2015). Helicobacter pylori and T Helper Cells: Mechanisms of Immune Escape and Tolerance. *Journal of Immunology Research, 2015*, 981328. doi:10.1155/2015/981328

Latz, E., Xiao, T. S., & Stutz, A. (2013). Activation and regulation of the inflammasomes. *Nat Rev Immunol, 13*(6), 397-411. doi:10.1038/nri3452

Lawingco, T., Chaudhury, S., Brookes, K. J., Guetta-Baranes, T., Guerreiro, R., Bras, J., . . . Morgan, K. (2021). Genetic variants in glutamate-, Aβ−, and tau-related pathways determine polygenic risk for Alzheimer's disease. *Neurobiology of Aging, 101*, 299.e213-299.e221. doi:https://doi.org/10.1016/j.neurobiolaging.2020.11.009

Lawlor, K. E., Feltham, R., Yabal, M., Conos, S. A., Chen, K. W., Ziehe, S., . . . Vince, J. E. (2017). XIAP Loss Triggers RIPK3- and Caspase-8-Driven IL-1β Activation and Cell Death as a Consequence of TLR-MyD88-Induced cIAP1-TRAF2 Degradation. *Cell Reports, 20*(3), 668-682. doi:10.1016/j.celrep.2017.06.073

Lee, S.-J., Nam, K.-I., Jin, H.-M., Cho, Y.-N., Lee, S.-E., Kim, T.-J., . . . Park, Y.-W. (2011). Bone destruction by receptor activator of nuclear factor κB ligand-expressing T cells in chronic gouty arthritis. *Arthritis Research & Therapy, 13*(5), R164. doi:10.1186/ar3483

Lee, Y. H., & Bae, S.-C. (2018). Correlation between circulating VEGF levels and disease activity in rheumatoid arthritis: a meta-analysis. *Zeitschrift für Rheumatologie, 77*(3), 240-248. doi:10.1007/s00393-016-0229-5

Lee, Y. H., & Song, G. G. (2012). Pathway analysis of genome-wide association studies on uric acid concentrations. *Hum Immunol, 73*(8), 805-810. doi:10.1016/j.humimm.2012.05.004

Leslie, R., O'Donnell, C. J., & Johnson, A. D. (2014). GRASP: analysis of genotype–phenotype results from 1390 genome-wide association studies and corresponding open access database. *Bioinformatics, 30*(12), i185-i194. doi:10.1093/bioinformatics/btu273

Li, C., Li, Z., Liu, S., Wang, C., Han, L., Cui, L., . . . Shi, Y. (2015). Genome-wide association analysis identifies three new risk loci for gout arthritis in Han Chinese. *Nat Commun, 6*, 7041. doi:10.1038/ncomms8041

Li, Q., Li, X., Kwong, J. S.-W., Chen, H., Sun, X., Tian, H., & Li, S. (2017). Diagnosis and treatment for hyperuricaemia and gout: a protocol for a systematic review of clinical practice guidelines and consensus statements. *BMJ Open, 7*.

Li, S., Sanna, S., Maschio, A., Busonero, F., Usala, G., Mulas, A., . . . Nagaraja, R. (2007). The GLUT9 Gene Is Associated with Serum Uric Acid Levels in Sardinia and Chianti Cohorts. *PLOS Genetics, 3*(11), e194. doi:10.1371/journal.pgen.0030194

Lieber, C. S., Jones, D. P., Losowsky, M. S., & Davidson, C. S. (1962). Interrelation of uric acid and ethanol metabolism in man. *The Journal of Clinical Investigation, 41*(10), 1863-1870. doi:10.1172/JCI104643

Lima Jr, H., Jacobson, L., Goldberg, M., Chandran, K., Diaz-Griffero, F., Lisanti, M. P., & Brojatsch, J. (2013). Role of lysosome rupture in controlling Nlrp3 signaling and necrotic cell death. *Cell Cycle, 12*(12), 1868-1878. doi:10.4161/cc.24903

LiotÉ, F., Champy, R., Moenner, M., Boval-Boizard, B., & Badet, J. (2003). Elevated angiogenin levels in synovial fluid from patients with inflammatory arthritis and secretion of angiogenin by cultured synovial fibroblasts. *Clinical & Experimental Immunology, 132*(1), 163-168. doi:https://doi.org/10.1046/j.1365-2249.2003.02117.x

Liu-Bryan, R. (2010). Intracellular innate immunity in gouty arthritis: role of NALP3 inflammasome. *Immunology and Cell Biology, 88*(1), 20-23. doi:10.1038/icb.2009.93

Liu-Bryan, R., Scott, P., Sydlaske, A., Rose David, M., & Terkeltaub, R. (2005). Innate immunity conferred by toll-like receptors 2 and 4 and myeloid differentiation factor 88 expression is pivotal to monosodium urate monohydrate crystal–induced inflammation. *Arthritis & Rheumatism, 52*(9), 2936-2946. doi:10.1002/art.21238

Liu, L., Xue, Y., Zhu, Y., Xuan, D., Yang, X., Liang, M., . . . Zou, H. (2016). Interleukin 37 limits monosodium urate crystal-induced innate immune responses in human and murine models of gout. *Arthritis Research & Therapy, 18*(1), 268. doi:10.1186/s13075-016-1167-y

Liu, Y., Yu, Q., & Chen, Y. (2018). Effect of silibinin on CFLAR-JNK pathway in oleic acid-treated HepG2 cells. *Biomedicine & Pharmacotherapy, 108*, 716-723. doi:https://doi.org/10.1016/j.biopha.2018.09.089

Liu, Y., Zhao, Q., Yin, Y., McNutt, M. A., Zhang, T., & Cao, Y. (2018). Serum levels of IL-17 are elevated in patients with acute gouty arthritis. *Biochemical and Biophysical Research Communications, 497*(3), 897-902. doi:https://doi.org/10.1016/j.bbrc.2018.02.166

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods, 25*(4), 402-408. doi:https://doi.org/10.1006/meth.2001.1262

Loeb, J. N. (1972). The influence of temperature on the solubility of monosodium urate. *Arthritis & Rheumatism, 15*(2), 189-192. doi:https://doi.org/10.1002/art.1780150209

Macgregor, S., Lind, P. A., Bucholz, K. K., Hansell, N. K., Madden, P. A., Richter, M. M., . . . Whitfield, J. B. (2009). Associations of ADH and ALDH2 gene variation with self report alcohol reactions, consumption and dependence: an integrated analysis. *Hum Mol Genet, 18*(3), 580-593. doi:10.1093/hmg/ddn372

Machiela, M. J., & Chanock, S. J. (2015). LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics, 31*(21), 3555-3557. doi:10.1093/bioinformatics/btv402

MacMullan, P., & McCarthy, G. (2012). Treatment and management of pseudogout: insights for the clinician. *Therapeutic Advances in Musculoskeletal Disease, 4*(2), 121-131. doi:10.1177/1759720X11432559

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), 341-353. doi:10.1038/s41584-018-0004-x

Male, D., Brostoff, J., Roth, D., & Roitt, I. (2012). *Immunology : With STUDENT CONSULT Online Access*. Edinburgh, UNITED KINGDOM: Elsevier.

Man Si, M., & Kanneganti, T. D. (2015). Regulation of inflammasome activation. *Immunological Reviews, 265*(1), 6-21. doi:10.1111/imr.12296

Mandal, A. K., & Mount, D. B. (2015). The Molecular Physiology of Uric Acid Homeostasis. *Annual Review of Physiology, 77*(1), 323-345. doi:10.1146/annurev-physiol-021113-170343

Manolio, T. A. (2010). Genomewide Association Studies and Assessment of the Risk of Disease. *New England Journal of Medicine, 363*(2), 166-176. doi:10.1056/NEJMra0905980

Marees, A. T., de Kluiver, H., Stringer, S., Vorspan, F., Curis, E., Marie-Claire, C., & Derks, E. M. (2018). A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. *International Journal of Methods in Psychiatric Research, 27*(2), e1608. doi:https://doi.org/10.1002/mpr.1608

Martillo, M. A., Nazzal, L., & Crittenden, D. B. (2014). The crystallization of monosodium urate. *Curr Rheumatol Rep, 16*(2), 400. doi:10.1007/s11926-013-0400-9

Martin William, J., Walton, M., & Harper, J. (2008). Resident macrophages initiating and driving inflammation in a monosodium urate monohydrate crystal–induced murine peritoneal model of acute gout. *Arthritis & Rheumatism, 60*(1), 281-289. doi:10.1002/art.24185

Mathes, E., O'Dea, E. L., Hoffmann, A., & Ghosh, G. (2008). NF-κB dictates the degradation pathway of IκBα. *The EMBO Journal, 27*(9), 1421-1421. doi:https://doi.org/10.1038/emboj.2008.91

Matsuo, H., Ichida, K., Takada, T., Nakayama, A., Nakashima, H., Nakamura, T., . . . Shinomiya, N. (2013). Common dysfunctional

variants in ABCG2 are a major cause of early-onset gout. *Sci Rep, 3*, 2014. doi:10.1038/srep02014

Matsuo, H., Nakayama, A., Sakiyama, M., Chiba, T., Shimizu, S., Kawamura, Y., . . . Shinomiya, N. (2014). ABCG2 dysfunction causes hyperuricemia due to both renal urate underexcretion and renal urate overload. *Scientific Reports, 4*, 3755. doi:10.1038/srep03755

https://www.nature.com/articles/srep03755#supplementary-information

Matsuo, H., Yamamoto, K., Nakaoka, H., Nakayama, A., Sakiyama, M., Chiba, T., . . . Shinomiya, N. (2016). Genome-wide association study of clinically defined gout identifies multiple risk loci and its association with clinical subtypes. *Ann Rheum Dis, 75*(4), 652-659. doi:10.1136/annrheumdis-2014-206191

Mavaddat, N., Michailidou, K., Dennis, J., Lush, M., Fachal, L., Lee, A., . . . Easton, D. F. (2019). Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *The American Journal of Human Genetics, 104*(1), 21-34. doi:10.1016/j.ajhg.2018.11.002

McCarty, D. J., & Hollander, J. L. (1961). IDENTIFICATION OF URATE CRYSTALS IN GOUTY SYNOVIAL FLUID. *Annals of Internal Medicine, 54*(3), 452-460. doi:10.7326/0003-4819-54-3-452

McGill, N. W., & Dieppe, P. A. (1991). Evidence for a promoter of urate crystal formation in gouty synovial fluid. *Annals of the Rheumatic Diseases, 50*(8), 558-561. doi:10.1136/ard.50.8.558

McKinney, C., Stamp, L. K., Dalbeth, N., Topless, R. K., Day, R. O., Kannangara, D. R., . . . Merriman, T. R. (2015). Multiplicative interaction of functional inflammasome genetic variants in determining the risk of gout. *Arthritis Res Ther, 17*, 288. doi:10.1186/s13075-015-0802-3

McQueen, F. M., Doyle, A., & Dalbeth, N. (2011). Imaging in gout - What can we learn from MRI, CT, DECT and US? *Arthritis Research & Therapy, 13*(6), 246. doi:10.1186/ar3489

Menu, P., & Vince, J. E. (2011). The NLRP3 inflammasome in health and disease: the good, the bad and the ugly. *Clinical & Experimental Immunology, 166*(1), 1-15. doi:10.1111/j.1365-2249.2011.04440.x

Merriman, T. R. (2015). An update on the genetic architecture of hyperuricemia and gout. *Arthritis Research & Therapy, 17*(1), 98. doi:10.1186/s13075-015-0609-2

Merriman, T. R., & Dalbeth, N. (2011). The genetic basis of hyperuricaemia and gout. *Joint Bone Spine, 78*(1), 35-40. doi:10.1016/j.jbspin.2010.02.027

Mikolajczak, S. A., Ma, B. Y., Yoshida, T., Yoshida, R., Kelvin, D. J., & Ochi, A. (2004). The Modulation of CD40 Ligand Signaling by Transmembrane CD28 Splice Variant in Human T Cells. *Journal of Experimental Medicine, 199*(7), 1025-1031. doi:10.1084/jem.20031705

Miranda-Aquino, T., Pérez-Topete, S. E., González-Padilla, C., Hernández-Del Río, J. E., Lomelí-Sánchez Ó, S., Esturau-Santaló, R. M., . . . González-Díaz, V. (2021). Asymptomatic hyperuricaemia and coronary artery disease. *Reumatol Clin, 17*(5), 263-267. doi:10.1016/j.reuma.2019.08.003

Myers, T. A., Chanock, S. J., & Machiela, M. J. (2020). LDlinkR: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations. *Frontiers in Genetics, 11*(157). doi:10.3389/fgene.2020.00157

Nakayama, A., Matsuo, H., Shimizu, T., Ogata, H., Takada, Y., Nakashima, H., . . . Shinomiya, N. (2013). A common missense variant of monocarboxylate transporter 9 (MCT9/SLC16A9) gene is associated with renal overload gout, but not with all gout susceptibility. *Human Cell, 26*(4), 133-136. doi:10.1007/s13577-013-0073-8

Nakayama, A., Nakaoka, H., Yamamoto, K., Sakiyama, M., Shaukat, A., Toyoda, Y., . . . Matsuo, H. (2017). GWAS of clinically defined gout and subtypes identifies multiple susceptibility loci that include

urate transporter genes. *Annals of the Rheumatic Diseases, 76*(5), 869.

Nardin, R. A., Fogerson, P. M., Nie, R., & Rutkove, S. B. (2010). Foot Temperature in Healthy Individuals: Effects of Ambient Temperature and Age. *Journal of the American Podiatric Medical Association, 100*(4), 258-264. doi:10.7547/1000258

Naredo, E., Uson, J., Jimenez-Palop, M., Martinez, A., Vicente, E., Brito, E., . . . Pascual, E. (2014). Ultrasound-detected musculoskeletal urate crystal deposition: which joints and what findings should be assessed for diagnosing gout? *Ann Rheum Dis, 73*(8), 1522-1528. doi:10.1136/annrheumdis-2013-203487

Nath, S. D., Voruganti, V. S., Arar, N. H., Thameem, F., Lopez-Alvarenga, J. C., Bauer, R., . . . Abboud, H. E. (2007). Genome Scan for Determinants of Serum Uric Acid Variability. *Journal of the American Society of Nephrology, 18*(12), 3156. doi:10.1681/ASN.2007040426

Negash, A. A., Olson, R. M., Griffin, S., & Gale, M., Jr. (2019). Modulation of calcium signaling pathway by hepatitis C virus core protein stimulates NLRP3 inflammasome activation. *PLOS Pathogens, 15*(2), e1007593. doi:10.1371/journal.ppat.1007593

Neogi, T., Jansen, T. L. T. A., Dalbeth, N., Fransen, J., Schumacher, H. R., Berendsen, D., . . . Taylor, W. J. (2015). 2015 Gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Annals of the Rheumatic Diseases, 74*, 1789-1798.

NICE. (2015). Chronic kidney disease in adults: assessment and management. *Clinical guideline 182.* Retrieved from https://www.nice.org.uk/guidance/cg182

Nyhan, W. L. (2005). Disorders of purine and pyrimidine metabolism. *Molecular Genetics and Metabolism, 86*(1), 25-33. doi:10.1016/j.ymgme.2005.07.027

Oda, M., Satta, Y., Takenaka, O., & Takahata, N. (2002). Loss of Urate Oxidase Activity in Hominoids and its Evolutionary Implications. *Molecular Biology and Evolution, 19*(5), 640-653. doi:10.1093/oxfordjournals.molbev.a004123

Ortiz-Bravo, E., Sieck, M. S., & Ralph Schumacher Jr, H. (1993). Changes in the proteins coating monosodium urate crystals during active and subsiding inflammation. immunogold studies of synovial fluid from patients with gout and of fluid obtained using the rat subcutaneous air pouch model. *Arthritis & Rheumatism, 36*(9), 1274-1285. doi:https://doi.org/10.1002/art.1780360912

Overgaard Nana, H., Jung, J.-W., Steptoe Raymond, J., & Wells James, W. (2014). CD4+/CD8+ double-positive T cells: more than just a developmental stage? *Journal of Leukocyte Biology, 97*(1), 31-38. doi:10.1189/jlb.1RU0814-382

Papin, S., Cuenin, S., Agostini, L., Martinon, F., Werner, S., Beer, H. D., . . . Tschopp, J. (2007). The SPRY domain of Pyrin, mutated in familial Mediterranean fever patients, interacts with inflammasome components and inhibits proIL-1β processing. *Cell Death & Differentiation, 14*(8), 1457-1466. doi:10.1038/sj.cdd.4402142

Parathithasan, N., Lee, W., Pianta, M., Oon, S., & Perera, W. (2016). Gouty arthropathy: Review of clinico-pathologic and imaging features. *J Med Imaging Radiat Oncol, 60*(1), 9-20. doi:10.1111/1754-9485.12356

Pascual, E., Batlle-Gualda, E., Martinez, A., Rosas, J., & Vela, P. (1999). Synovial fluid analysis for diagnosis of intercritical gout. *Annals of Internal Medicine, 131*(10), 756-759. doi:10.7326/0003-4819-131-10-199911160-00007

Pattaro, C., Teumer, A., Gorski, M., Chu, A. Y., Li, M., Mijatovic, V., . . . Consortium, I. (2016). Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nature Communications, 7*(1), 10023. doi:10.1038/ncomms10023

Paul, B. J., Anoopkumar, K., & Krishnan, V. (2017). Asymptomatic hyperuricemia: is it time to intervene? *Clinical Rheumatology, 36*(12), 2637-2644. doi:10.1007/s10067-017-3851-y

Paul, H., Reginato, A. J., & Schumacher, H. R. (1983). Morphological characteristics of monosodium urate: a transmission electron microscopic study of intact natural and synthetic crystals. *Annals of the Rheumatic Diseases, 42*(1), 75. doi:10.1136/ard.42.1.75

Peláez-Ballestas, I., Hernández Cuevas, C., Burgos-Vargas, R., Hernández Roque, L., Terán, L., Espinoza, J., . . . Vázquez-Mellado, J. (2010). Diagnosis of Chronic Gout: Evaluating the American College of Rheumatology Proposal, European League Against Rheumatism Recommendations, and Clinical Judgment. *The Journal of Rheumatology, 37*(8), 1743. doi:10.3899/jrheum.091385

Perez-Ruiz, F., & Dalbeth, N. (2019). Gout. *Rheumatic Disease Clinics, 45*(4), 583-591. doi:10.1016/j.rdc.2019.08.001

Perl, K., Ushakov, K., Pozniak, Y., Yizhar-Barnea, O., Bhonker, Y., Shivatzki, S., . . . Shamir, R. (2017). Reduced changes in protein compared to mRNA levels across non-proliferating tissues. *BMC Genomics, 18*(1), 305. doi:10.1186/s12864-017-3683-9

Petrilli, V., Papin, S., Dostert, C., Mayor, A., Martinon, F., & Tschopp, J. (2007). Activation of the NALP3 inflammasome is triggered by low intracellular potassium concentration. *Cell Death Differ, 14*(9), 1583-1589. doi:10.1038/sj.cdd.4402195

Phipps-Green, A. J., Merriman, M. E., Topless, R., Altaf, S., Montgomery, G. W., Franklin, C., . . . Merriman, T. R. (2016). Twenty-eight loci that influence serum urate levels: analysis of association with gout. *Ann Rheum Dis, 75*(1), 124-130. doi:10.1136/annrheumdis-2014-205877

Pineda, C., Amezcua-Guerra, L. M., Solano, C., Rodriguez-Henríquez, P., Hernández-Díaz, C., Vargas, A., . . . Gutiérrez, M. (2011). Joint and tendon subclinical involvement suggestive of gouty arthritis in

asymptomatic hyperuricemia: an ultrasound controlled study. *Arthritis Research & Therapy, 13*(1), R4. doi:10.1186/ar3223

Polimanti, R., & Gelernter, J. (2018). ADH1B: From alcoholism, natural selection, and cancer to the human phenome. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics, 177*(2), 113-125. doi:10.1002/ajmg.b.32523

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., . . . Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics, 81*(3), 559-575. doi:10.1086/519795

Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., . . . Harvard. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature, 460*(7256), 748-752. doi:10.1038/nature08185

Qing, Y. F., Zhou, J. G., Zhang, Q. B., Wang, D. S., Li, M., Yang, Q. B., . . . Zeng, M. (2013). Association of TLR4 Gene rs2149356 polymorphism with primary gouty arthritis in a case-control study. *PLoS One, 8*(5), e64845. doi:10.1371/journal.pone.0064845

Qu, Y., Misaghi, S., Newton, K., Maltzman, A., Izrael-Tomasevic, A., Arnott, D., & Dixit, V. M. (2016). NLRP3 recruitment by NLRC4 during Salmonella infection. *Journal of Experimental Medicine, 213*(6), 877-885. doi:10.1084/jem.20132234

Ragab, G., Elshahaly, M., & Bardin, T. (2017). Gout: An old disease in new perspective - A review. *J Adv Res, 8*(5), 495-511. doi:10.1016/j.jare.2017.04.008

Rathinam, V. A., Vanaja, S. K., & Fitzgerald, K. A. (2012). Regulation of inflammasome signaling. *Nat Immunol, 13*(4), 333-342. doi:10.1038/ni.2237

Razmara, M., Srinivasula, S. M., Wang, L., Poyet, J. L., Geddes, B. J., DiStefano, P. S., . . . Alnemri, E. S. (2002). CARD-8 protein, a new CARD family member that regulates caspase-1 activation and

apoptosis. *J Biol Chem, 277*(16), 13952-13958. doi:10.1074/jbc.M107811200

Riches, P. L., Wright, A. F., & Ralston, S. H. (2009). Recent insights into the pathogenesis of hyperuricaemia and gout. *Human Molecular Genetics, 18*(R2), R177-R184. doi:10.1093/hmg/ddp369

Richette, P., Doherty, M., Pascual, E., Barskova, V., Becce, F., Castañeda-Sanabria, J., . . . Bardin, T. (2017). 2016 updated EULAR evidence-based recommendations for the management of gout. *Annals of the Rheumatic Diseases, 76*, 29-42.

Richette, P., Doherty, M., Pascual, E., Barskova, V., Becce, F., Castaneda, J., . . . Bardin, T. (2019). 2018 updated European League Against Rheumatism evidence-based recommendations for the diagnosis of gout. *Annals of the Rheumatic Diseases, 79*(1), 31. doi:10.1136/annrheumdis-2019-215315

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

Rokhlin, V., Szlam, A., & Tygert, M. (2009). A Randomized Algorithm for Principal Component Analysis. *SIAM Journal on Matrix Analysis and Applications, 31*(3), 1100-1124. doi:10.1137/080736417

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

Rothschild, B. M., Tanke, D., & Carpenter, K. (1997). Tyrannosaurs suffered from gout. *Nature, 387*(6631), 357-357. doi:10.1038/387357a0

Rowe, P. S. N. (2004). The wrickkened pathways of FGF23, MEPE and PHEX. *Critical reviews in oral biology and medicine : an official publication of the American Association of Oral Biologists, 15*(5), 264-281. doi:10.1177/154411130401500503

Ruoff, G., & Edwards, N. L. (2016). Overview of Serum Uric Acid Treatment Targets in Gout: Why Less Than 6 mg/dL? *Postgraduate Medicine, 128*(7), 706-715. doi:10.1080/00325481.2016.1221732

Sakiyama, M., Matsuo, H., Akashi, A., Shimizu, S., Higashino, T., Kawaguchi, M., . . . Shinomiya, N. (2017). Independent effects of ADH1B and ALDH2 common dysfunctional variants on gout risk. *Sci Rep, 7*(1), 2500. doi:10.1038/s41598-017-02528-z

Sakiyama, M., Matsuo, H., Nakaoka, H., Yamamoto, K., Nakayama, A., Nakamura, T., . . . Shinomiya, N. (2016). Identification of rs671, a common variant of ALDH2, as a gout susceptibility locus. *Scientific Reports, 6*, 25360. doi:10.1038/srep25360

https://www.nature.com/articles/srep25360#supplementary-information

Sakiyama, M., Matsuo, H., Shimizu, S., Nakashima, H., Nakamura, T., Nakayama, A., . . . Shinomiya, N. (2016). The effects of URAT1/SLC22A12 nonfunctional variants,R90H and W258X, on serum uric acid levels and gout/hyperuricemia progression. *Scientific Reports, 6*(1), 20148. doi:10.1038/srep20148

Sakiyama, M., Matsuo, H., Shimizu, S., Nakashima, H., Nakayama, A., Chiba, T., . . . Shinomiya, N. (2014). A Common Variant of Organic Anion Transporter 4 OAT4/SLC22A11) Gene Is Associated with Renal Underexcretion Type Gout. *Drug Metabolism and Pharmacokinetics, 29*(2), 208-210. doi:10.2133/dmpk.DMPK-13-NT-070

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 (Oxford), 51*(11), 2004-2012. doi:10.1093/rheumatology/kes183

Scanu, A., Luisetto, R., Oliviero, F., Gruaz, L., Sfriso, P., Burger, D., & Punzi, L. (2015). High-density lipoproteins inhibit urate crystal-

induced inflammation in mice. *Ann Rheum Dis, 74*(3), 587-594. doi:10.1136/annrheumdis-2013-203803

Scanu, A., Oliviero, F., Gruaz, L., Sfriso, P., Pozzuoli, A., Frezzato, F., . . . Punzi, L. (2010). High-density lipoproteins downregulate CCL2 production in human fibroblast-like synoviocytes stimulated by urate crystals. *Arthritis Research & Therapy, 12*(1), R23. doi:10.1186/ar2930

Scherer, H. U., & Burmester, G.-R. (2011). Adaptive immunity in rheumatic diseases – Bystander or pathogenic player? *Best Practice & Research Clinical Rheumatology, 25*(6), 785-800. doi:https://doi.org/10.1016/j.berh.2011.11.005

Schett, G., Schauer, C., Hoffmann, M., & Herrmann, M. (2015). Why does the gout attack stop? A roadmap for the immune pathogenesis of gout. *RMD Open, 1*(Suppl 1), e000046. doi:10.1136/rmdopen-2015-000046

Schlesinger, N. (2004). Management of Acute and Chronic Gouty Arthritis. *Drugs, 64*(21), 2399-2416. doi:10.2165/00003495-200464210-00003

Schlesinger, N., Alten, R. E., Bardin, T., Schumacher, H. R., Bloch, M., Gimona, A., . . . So, A. K. (2012). Canakinumab for acute gouty arthritis in patients with limited treatment options: results from two randomised, multicentre, active-controlled, double-blind trials and their initial extensions. *Annals of the Rheumatic Diseases, 71*(11), 1839.

Schorn, C., Janko, C., Krenn, V., Zhao, Y., Munoz, L. E., Schett, G., & Herrmann, M. (2012). Bonding the foe - NETting neutrophils immobilize the pro-inflammatory monosodium urate crystals. *Front Immunol, 3*, 376. doi:10.3389/fimmu.2012.00376

Scott, P., Ma, H., Viriyakosol, S., Terkeltaub, R., & Liu-Bryan, R. (2006). Engagement of CD14 Mediates the Inflammatory Potential of Monosodium Urate Crystals. *The Journal of Immunology, 177*(9), 6370-6378. doi:10.4049/jimmunol.177.9.6370

Scuteri, A., Sanna, S., Chen, W.-M., Uda, M., Albai, G., Strait, J., . . . Abecasis, G. R. (2007). Genome-Wide Association Scan Shows Genetic Variants in the FTO Gene Are Associated with Obesity-Related Traits. *PLOS Genetics, 3*(7), e115. doi:10.1371/journal.pgen.0030115

Shi, J., Levinson, D. F., Duan, J., Sanders, A. R., Zheng, Y., Pe'er, I., . . . Gejman, P. V. (2009). Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature, 460*(7256), 753-757. doi:10.1038/nature08192

Shi, Y., Mucsi Ashley, D., & Ng, G. (2009). Monosodium urate crystals in inflammation and immunity. *Immunological Reviews, 233*(1), 203-217. doi:10.1111/j.0105-2896.2009.00851.x

Silverman, W., Locovei, S., & Dahl, G. (2008). Probenecid, a gout remedy, inhibits pannexin 1 channels. *American Journal of Physiology-Cell Physiology, 295*(3), C761-C767. doi:10.1152/ajpcell.00227.2008

Smemo, S., Tena, J. J., Kim, K.-H., Gamazon, E. R., Sakabe, N. J., Gómez-Marín, C., . . . Nóbrega, M. A. (2014). Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature, 507*(7492), 371-375. doi:10.1038/nature13138

Stamp, L., & Dalbeth, N. (2017). Urate-lowering therapy for asymptomatic hyperuricaemia: A need for caution. *Seminars in Arthritis and Rheumatism, 46*(4), 457-464. doi:https://doi.org/10.1016/j.semarthrit.2016.07.015

Stewart, S., Dalbeth, N., Vandal, A. C., Allen, B., Miranda, R., & Rome, K. (2017). Are ultrasound features at the first metatarsophalangeal joint associated with clinically-assessed pain and function? A study of people with gout, asymptomatic hyperuricaemia and normouricaemia. *Journal of Foot and Ankle Research, 10*(1), 22. doi:10.1186/s13047-017-0203-8

Storck, A. (1764). *An essay on the use and effects of the root of the Colchicum autumnale, or Meadow Saffron.* London: Becket, T. de Honet, PA.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., . . . Collins, R. (2015). UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine, 12*(3), e1001779. doi:10.1371/journal.pmed.1001779

Sumpter, N., Merriman, T., Reynolds, R., Abhishek, A., Andres, M., Dalbeth, N., . . . Uhlig, T. (2019). *Association of a gout polygenic risk score with disease severity phenotypes amongst Caucasian gout patients in three independent cohorts [Abstract]*. Paper presented at the American College of Rheumatology 2019. https://acrabstracts.org/abstract/association-of-a-gout-polygenic-risk-score-with-disease-severity-phenotypes-amongst-caucasian-gout-patients-in-three-independent-cohorts

Sutterwala, F. S., Haasken, S., & Cassel, S. L. (2014). Mechanism of NLRP3 inflammasome activation. *Ann N Y Acad Sci, 1319*, 82-95. doi:10.1111/nyas.12458

Sydenham, T. (1683). *Tractus de Podagra et Hydrope* (G. Kettibly Ed.). London.

SydenhamSociety. (1891). *The Genuine Works of Hippocrates* (F. Adams Ed.). London.

Tan, B. M., Zammit, N. W., Yam, A. O., Slattery, R., Walters, S. N., Malle, E., & Grey, S. T. (2013). Baculoviral inhibitors of apoptosis repeat containing (BIRC) proteins fine-tune TNF-induced nuclear factor κB and c-Jun N-terminal kinase signalling in mouse pancreatic beta cells. *Diabetologia, 56*(3), 520-532. doi:10.1007/s00125-012-2784-x

Terkeltaub, R. (2017). What makes gouty inflammation so variable? *BMC Medicine, 15*(1), 158. doi:10.1186/s12916-017-0922-5

Terkeltaub, R., Tenner Andrea, J., Kozin, F., & Ginsberg Mark, H. (1983). Plasma Protein Binding by Monosodium Urate Crystals. *Arthritis & Rheumatism, 26*(6), 775-783. doi:10.1002/art.1780260612

Terkeltaub, R. A., Dyer, C. A., Martin, J., & Curtiss, L. K. (1991). Apolipoprotein (apo) E inhibits the capacity of monosodium urate

crystals to stimulate neutrophils. Characterization of intraarticular apo E and demonstration of apo E binding to urate crystals in vivo. *The Journal of Clinical Investigation, 87*(1), 20-26. doi:10.1172/JCI114971

Thiele, R. G., & Schlesinger, N. (2007). Diagnosis of gout by ultrasound. *Rheumatology, 46*(7), 1116-1121. doi:10.1093/rheumatology/kem058

Tin, A., Marten, J., Halperin Kuhns, V. L., Li, Y., Wuttke, M., Kirsten, H., . . . Program, V. A. M. V. (2019). Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels. *Nature Genetics, 51*(10), 1459-1474. doi:10.1038/s41588-019-0504-x

Tin, A., Woodward, O. M., Kao, W. H., Liu, C. T., Lu, X., Nalls, M. A., . . . Consortia, C. (2011). Genome-wide association study for serum urate concentrations and gout among African Americans identifies genomic risk loci and a novel URAT1 loss-of-function allele. *Hum Mol Genet, 20*(20), 4056-4068. doi:10.1093/hmg/ddr307

Torres, R. J., & Puig, J. G. (2007). Hypoxanthine-guanine phosophoribosyltransferase (HPRT) deficiency: Lesch-Nyhan syndrome. *Orphanet Journal of Rare Diseases, 2*(1), 48. doi:10.1186/1750-1172-2-48

Tschopp, J., & Schroder, K. (2010). NLRP3 inflammasome activation: The convergence of multiple signalling pathways on ROS production? *Nat Rev Immunol, 10*(3), 210-215. doi:10.1038/nri2725

Tseng, C.-C., Liao, W.-T., Wong, M.-C., Chen, C.-J., Lee, S.-C., Yen, J.-H., & Chang, S.-J. (2021). Cell lineage-specific methylome and genome alterations in gout. *Aging, 13*(3), 3843-3865. doi:10.18632/aging.202353

Tuladhar, S., & Kanneganti, T.-D. (2020). NLRP12 in innate immunity and inflammation. *Molecular Aspects of Medicine, 76*, 100887. doi:https://doi.org/10.1016/j.mam.2020.100887

Turner, S., Armstrong, L. L., Bradford, Y., Carlson, C. S., Crawford, D. C., Crenshaw, A. T., . . . Ritchie, M. D. (2011). Quality Control Procedures for Genome-Wide Association Studies. *Current Protocols in Human Genetics, 68*(1), 1.19.11-11.19.18. doi:https://doi.org/10.1002/0471142905.hg0119s68

Udden, S. M. N., Kwak, Y.-T., Godfrey, V., Khan, M. A. W., Khan, S., Loof, N., . . . Zaki, H. (2019). NLRP12 suppresses hepatocellular carcinoma via downregulation of cJun N-terminal kinase activation in the hepatocyte. *eLife, 8*, e40396. doi:10.7554/eLife.40396

van der Wijst, M. G. P., de Vries, D. H., Brugge, H., Westra, H.-J., & Franke, L. (2018). An integrative approach for building personalized gene regulatory networks for precision medicine. *Genome Medicine, 10*(1), 96. doi:10.1186/s13073-018-0608-4

Vargas-Santos, A. B., Taylor, W. J., & Neogi, T. (2016). Gout Classification Criteria: Update and Implications. *Curr Rheumatol Rep, 18*(1). doi:doi:10.1007/s11926-016-0594-8

Venuprasad, K., & Theiss, A. L. (2021). NLRP6 in host defense and intestinal inflammation. *Cell Reports, 35*(4). doi:10.1016/j.celrep.2021.109043

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics, 101*(1), 5-22. doi:10.1016/j.ajhg.2017.06.005

Vitart, V., Rudan, I., Hayward, C., Gray, N. K., Floyd, J., Palmer, C. N. A., . . . Wright, A. F. (2008). SLC2A9 is a newly identified urate transporter influencing serum urate concentration, urate excretion and gout. *Nature Genetics, 40*, 437. doi:10.1038/ng.106

https://www.nature.com/articles/ng.106#supplementary-information

Vladimer, Gregory I., Weng, D., Paquette, Sara W. M., Vanaja, Sivapriya K., Rathinam, Vijay A. K., Aune, Marie H., . . . Lien, E. (2012). The NLRP12 Inflammasome Recognizes Yersinia pestis. *Immunity, 37*(1), 96-107. doi:10.1016/j.immuni.2012.07.006

Voruganti, V. S., Kent, J. W., Jr., Debnath, S., Cole, S. A., Haack, K., Goring, H. H., . . . Comuzzie, A. G. (2013). Genome-wide association analysis confirms and extends the association of SLC2A9 with serum uric acid levels to Mexican Americans. *Front Genet, 4*, 279. doi:10.3389/fgene.2013.00279

Wain, L. V., Shrine, N., Miller, S., Jackson, V. E., Ntalla, I., Artigas, M. S., . . . Hall, I. P. (2015). Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *The Lancet Respiratory Medicine, 3*(10), 769-781. doi:10.1016/S2213-2600(15)00283-0

Wallace, C., Newhouse, S. J., Braund, P., Zhang, F., Tobin, M., Falchi, M., . . . Munroe, P. B. (2008). Genome-wide Association Study Identifies Genes for Biomarkers of Cardiovascular Disease: Serum Urate and Dyslipidemia. *The American Journal of Human Genetics, 82*(1), 139-149. doi:10.1016/j.ajhg.2007.11.001

Wallace, K. L., Riedel, A. A., Joseph-Ridge, N., & Wortmann, R. (2004). Increasing prevalence of gout and hyperuricemia over 10 years among older adults in a managed care population. *The Journal of Rheumatology, 31*, 1582-1587.

Wallace, S. L. (1964). A short history of the gout and the rheumatic diseases. W. S. C. Copeman, M.D., F.R.C.P. Berkeley and Los Angeles, University of California Press, 1964, 236 pp. *Arthritis & Rheumatism, 7*(6), 722-723. doi:https://doi.org/10.1002/art.1780070613

Wang, J., Liu, S., Wang, B., Miao, Z., Han, L., Chu, N., . . . Ma, X. (2012). Association between gout and polymorphisms in GCKR in male Han Chinese. *Human Genetics, 131*(7), 1261-1265. doi:10.1007/s00439-012-1151-9

Wang, Q., Guo, L.-H., Li, X.-L., Zhao, C.-K., Li, M.-X., Wang, L., . . . Xu, H.-X. (2018). Differentiating the acute phase of gout from the intercritical phase with ultrasound and quantitative shear wave

elastography. *European Radiology, 28*(12), 5316-5327. doi:10.1007/s00330-018-5529-5

Watanabe, K., Taskesen, E., van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nature Communications, 8*(1), 1826. doi:10.1038/s41467-017-01261-5

Wawrocki, S., & Druszczynska, M. (2017). Inflammasomes in Mycobacterium tuberculosis-Driven Immunity. *The Canadian journal of infectious diseases & medical microbiology = Journal canadien des maladies infectieuses et de la microbiologie medicale, 2017*, 2309478-2309478. doi:10.1155/2017/2309478

Wei, C., Li, J., & Bumgarner, R. E. (2004). Sample size for detecting differentially expressed genes in microarray experiments. *BMC Genomics, 5*, 87-87. doi:10.1186/1471-2164-5-87

Welsh, S. (2017). Genotyping of 500,000 UK Biobank participants: Description of sample processing workflow and preparation of DNA for genotyping.

Welsh, S., Peakman, T., Sheard, S., & Almond, R. (2017). Comparison of DNA quantification methodology used in the DNA extraction protocol for the UK Biobank cohort. *BMC Genomics, 18*(1), 26. doi:10.1186/s12864-016-3391-x

Wen, C. C., Yee, S. W., Liang, X., Hoffmann, T. J., Kvale, M. N., Banda, Y., . . . Giacomini, K. M. (2015). Genome-wide association study identifies ABCG2 (BCRP) as an allopurinol transporter and a determinant of drug response. *Clin Pharmacol Ther, 97*(5), 518-525. doi:10.1002/cpt.89

Wigginton, J. E., Cutler, D. J., & Abecasis, G. R. (2005). A Note on Exact Tests of Hardy-Weinberg Equilibrium. *The American Journal of Human Genetics, 76*(5), 887-893. doi:https://doi.org/10.1086/429864

Wilcox, W. R., & Khalaf, A. A. (1975). Nucleation of monosodium urate crystals. *Annals of the Rheumatic Diseases, 34*(4), 332-339. doi:10.1136/ard.34.4.332

Williams, P. T. (2008). Effects of diet, physical activity and performance, and body weight on incident gout in ostensibly healthy, vigorously active men. *The American Journal of Clinical Nutrition, 87*(5), 1480-1487. doi:10.1093/ajcn/87.5.1480

Woodward, O. M., Köttgen, A., Coresh, J., Boerwinkle, E., Guggino, W. B., & Köttgen, M. (2009). Identification of a urate transporter, ABCG2, with a common functional polymorphism causing gout. *Proceedings of the National Academy of Sciences, 106*(25), 10338.

Wrigley, R., Phipps-Green, A. J., Topless, R. K., Major, T. J., Cadzow, M., Riches, P., . . . Merriman, T. R. (2020). Pleiotropic effect of the ABCG2 gene in gout: involvement in serum urate levels and progression from hyperuricemia to gout. *Arthritis Research & Therapy, 22*(1), 45. doi:10.1186/s13075-020-2136-z

Wu, X., D., M., Lee, C. C., & Caskey, C. T. (1992). Two independent mutational events in the loss of urate oxidase during hominoid evolution. *Journal of Molecular Evolution, 34*(1), 78-84. doi:https://doi.org/10.1007/BF00163854

Xia, X., Jin, J., Chen, Z.-J., Zhou, Z., Chen, H., Zhang, C., . . . Sun, L. (2020). Unraveling the genetic causes in large pedigrees with gout by whole-exome sequencing. *International journal of molecular medicine, 45*(4), 1047-1058. doi:10.3892/ijmm.2020.4501

Xiao, J., Zhang, X. L., Fu, C., Han, R., Chen, W., Lu, Y., & Ye, Z. (2015). Soluble uric acid increases NALP3 inflammasome and interleukin-1beta expression in human primary renal proximal tubule epithelial cells through the Toll-like receptor 4-mediated pathway. *Int J Mol Med, 35*(5), 1347-1354. doi:10.3892/ijmm.2015.2148

Xiaohong, W., Jun, Z., Hongmei, G., & Fan, Q. (2019). CFLAR is a critical regulator of cerebral ischaemia-reperfusion injury through regulating inflammation and endoplasmic reticulum (ER) stress. *Biomedicine & Pharmacotherapy, 117*, 109155. doi:https://doi.org/10.1016/j.biopha.2019.109155

Yabal, M., Müller, N., Adler, H., Knies, N., Groß, Christina J., Damgaard, Rune B., . . . Jost, Philipp J. (2014). XIAP Restricts TNF- and RIP3-Dependent Cell Death and Inflammasome Activation. *Cell Reports, 7*(6), 1796-1808. doi:10.1016/j.celrep.2014.05.008

Yang, Q., Köttgen, A., Dehghan, A., Smith, A. V., Glazer, N. L., Chen, M.-H., . . . Coresh, J. (2010). Multiple Genetic Loci Influence Serum Urate Levels and Their Relationship With Gout and Cardiovascular Disease Risk Factors. *Circulation: Cardiovascular Genetics, 3*(6), 523.

Yokoyama, A., Yokoyama, T., Mizukami, T., Matsui, T., Kimura, M., Matsushita, S., . . . Maruyama, K. (2016). Alcohol Dehydrogenase-1B (rs1229984) and Aldehyde Dehydrogenase-2 (rs671) Genotypes and Alcoholic Ketosis Are Associated with the Serum Uric Acid Level in Japanese Alcoholic Men. *Alcohol Alcohol, 51*(3), 268-274. doi:10.1093/alcalc/agv123

Yu, J., Riou, C., Davidson, D., Minhas, R., Robson Jeffrey, D., Julius, M., . . . Veillette, A. (2001). Synergistic Regulation of Immunoreceptor Signaling by SLP-76-Related Adaptor Clnk and Serine/Threonine Protein Kinase HPK-1. *Molecular and Cellular Biology, 21*(18), 6102-6112. doi:10.1128/MCB.21.18.6102-6112.2001

Yu, K.-H., Chang, P.-Y., Chang, S.-C., Wu-Chou, Y.-H., Wu, L.-A., Chen, D.-P., . . . Lu, J.-J. (2017). A comprehensive analysis of the association of common variants of ABCG2 with gout. *Scientific Reports, 7*(1), 9988. doi:10.1038/s41598-017-10196-2

Yu, Y., Yang, J., Fu, S., Xue, Y., Liang, M., Xuan, D., . . . Zou, H. (2019). Leptin Promotes Monosodium Urate Crystal–Induced Inflammation in Human and Murine Models of Gout. *The Journal of Immunology*, ji1801097. doi:10.4049/jimmunol.1801097

Zelenchuk, L. V., Hedge, A.-M., & Rowe, P. S. N. (2015). Age dependent regulation of bone-mass and renal function by the MEPE ASARM-motif. *Bone, 79*, 131-142. doi:10.1016/j.bone.2015.05.030

Zhang, Y., Chen, C., Choi, H., Chaisson, C., Hunter, D., Niu, J., & Neogi, T. (2012). Purine-rich foods intake and recurrent gout attacks. *Annals of the Rheumatic Diseases, 71*(9), 1448.

Zhang, Y., & Lee, M. T. M. (2021). Serum Urate Polygenic Risk Score Can Improve Gout Risk Prediction: A Large-Scale Cohort Study. *Frontiers in Genetics, 11*(1858). doi:10.3389/fgene.2020.604219

Zhang, Y., Liu, K., Ma, L., Liu, K., Shi, X., Zhang, Y., . . . Jin, T. (2016). Associations of gout with polymorphisms in SLC2A9, WDR1, CLNK, PKD2, and ABCG2 in Chinese Han and Tibetan populations. *Int J Clin Exp Pathol [Internet], 9*, 7503-7517.

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. *The American Journal of Medicine, 119*(9), 800.e811-800.e816. doi:https://doi.org/10.1016/j.amjmed.2006.01.020

Zhao, Y., Yang, J., Shi, J., Gong, Y.-N., Lu, Q., Xu, H., . . . Shao, F. (2011). The NLRC4 inflammasome receptors for bacterial flagellin and type III secretion apparatus. *Nature, 477*(7366), 596-600. doi:10.1038/nature10510

Zhou, R., Tardivel, A., Thorens, B., Choi, I., & Tschopp, J. (2009). Thioredoxin-interacting protein links oxidative stress to inflammasome activation. *Nature Immunology, 11*, 136. doi:10.1038/ni.1831

https://www.nature.com/articles/ni.1831#supplementary-information

Zhu, Y., Pandya, B. J., & Choi, H. K. (2011). Prevalence of gout and hyperuricemia in the US general population: the National Health and Nutrition Examination Survey 2007-2008. *Arthritis Rheum, 63*(10), 3136-3141. doi:10.1002/art.30520

Appendix 1. UK Biobank application (ID project 45987)

Application details

The purpose of the application is for UK Biobank to determine whether the proposed research project is health-related, feasible and in the public interest. For this, we require a brief synopsis of the research plan (i.e. a description of the aims, methods and intended outputs) rather than a full scientific review. Please refer to the online help for guidance and examples.

A1. Project title (200 characters): Development of a composite genetic, comorbid, and lifestyle factor related risk prediction model for gout, and cardiovascular diseases in people with gout.

A2. Research question(s) and aim(s) (up to 5000 characters or 200 words):

Hyperuricaemia is the main identified risk factor for gout. This is caused by either overproduction or impaired renal and gastrointestinal excretion of urate. Once serum urate exceeds the limit of solubilisation, it precipitates and forms crystals in and around the joints. Research suggests that a quarter of people with hyperuricaemia have asymptomatic monosodium urate deposition, while, only 10% people with hyperuricaemia develop symptomatic gout.

While the molecular mechanisms and genetic variants associated with serum urate concentration are well understood, the genetic factors contributing to progression from asymptomatic crystal deposition to gout remain unclear. Furthermore, gout associates with other metabolic and cardiovascular diseases that worsens quality of life and survival.

Therefore, this research aims to generate polygenic risk score (PRS) models with and without additional lifestyle, and comorbid factors for gout and its associated comorbidities. The specific objectives of this study are to develop a PRS for differentiating people with:

[1] gout from general population,

[2] gout from hyperuricaemic controls without gout,

[3] gout and cardiovascular comorbidities from people with gout without cardiovascular comorbidities.

PRS have the potential to identify novel genetic variants or pathways associated with a particular phenotype, and, may result in further basic science research.

The majority of applications to UK Biobank are for data only. As such, the first two questions we ask are whether your application involves access to samples or re-contact as this will require some additional information and as is set out in the Access Procedures (our data are not depletable, but our samples and re-contact opportunities are depletable) recontact/sample applications are assessed to a different (more exacting) standard.

Does your project require biological samples?

☐ Yes

☑ No

Does your project require UK Biobank to re-contact participants:

☐ Yes

☑ No

Please provide information on each of the following:

A3. The background and scientific rationale of the proposed research project in general (up to 5000 characters or 300 words):

Gout is the most common form of inflammatory arthritis, and affects 2.5% adults in the UK. It is caused by hyperuricaemia, and, once the serum urate exceeds a certain threshold, it deposits inside and around the joints as monosodium urate (MSU) crystals. Once released into the joints, the MSU crystals trigger an immune response by activating the NLRP3 inflammasome and Toll-like receptor pathway that together promote the synthesis of pro-inflammatory cytokines e.g. IL-1β. Although it is a treatable condition, gout's management remains suboptimal and patients often progress to a stage characterised by frequent flares, tophi, and development of other disorders, such as diabetes, obesity, hypertension, hyperlipidaemia, heart attacks, and renal failure.

Even though the heritability of gout is estimated at 40-70%, GWAS have allowed the identification of genetic variants that explain only 7% of that heritability, with most of the identified genetic variants associated with serum urate levels. What is more intriguing, despite being crucial to the pathogenesis of gout, hyperuricemia does not always causes gout. In fact, only 10-12% of the hyperuricaemic individuals develop gout. Similarly, approximately 60% people with gout have high blood pressure, while 5-10% have ischaemic heart disease or its complications, and, currently there is no predictive model for identifying these people at high risk of cardiovascular comorbidities.

Though GWAS have been a powerful tool for the identification of novel pathways or polymorphisms, they are still limited, since they only identify common variants and those exerting small effects are not detected, leaving most of the missing heritability unexplained. PRS have emerged as an alternative not only to address missing heritability in complex diseases by tracking back to the variants contributing to the scores, but also to generate models that could help classify individuals as cases or controls based on their genetic profiles.

A4. A brief description of the method(s) to be used (up to 5000 characters or 300 words):

Three PRS models will be generated as outlined above. Cardiovascular comorbidities will include hypertension, hyperlipidaemia, and ischaemic heart disease respectively. We will randomly select upto four age (+/- 5 years), sex matched controls for each case in the UK Biobank. We will do further modelling including body mass index (BMI), deprivation score; current smoking status, alcohol consumption, diuretic prescription along with the PRS for gout. We will also include diabetes, hypertension and hyperlipidaemia as additional covariates in the PRS for angina and myocardial infarction.

PRS will be calculated using a software package that computes scores for each individual from raw genotype data, at the best fit P-value threshold based on GWAS summary statistics. For this purpose, the target datasets will be assembled from raw genotype information for gout cases (or asymptomatic hyperuricemia cases - once biomarkers data become available) and controls. The training or base data sets will consist on publicly available summary statistics from GWAS for gout, and each cardiovascular disease. In the training dataset, we will order variants according to the p value of association with

phenotype, and select variants that associate with the phenotype at < user defined p value. Several p value thresholds will be tested to identify the variants that provides the highest goodness of fit assessed using R^2 characteristic for the phenotype in question. The PRS will be calculated for each individual in the target dataset, based on alleles present at each selected variant, and weighed according to the β-statistic in the base dataset. The distribution of the scores will be analysed, and regression analyses will be conducted to test the association between the scores and the phenotypes. Finally, ROC curves will be generated to evaluate predictive models. Logistic regression will be performed to undertake additional modelling including lifestyle and comorbid conditions.

A5. The type and size of dataset required (e.g., case-control subset, men only, imaging data only, whole cohort, etc.) (up to 5000 characters or 100 words):

We require data on genotype, age, sex, deprivation index, BMI, alcohol consumption, smoking status, comorbidities, prescriptions, serum urate, and lipid profile from the first visit from the entire UK biobank. We request access to data for the whole cohort in order to select cases and matched controls. Diagnosis and urate lowering treatment prescription will allow ascertainment of gout status. Self-reported comorbidity data will allow ascertainment of comorbidities.

Since one objective involves individuals with hyperuricemia without gout, we need serum urate to select cases and controls. Therefore, this will be a phased application to access biomarker data once available.

A6. The expected value of the research (taking into account the public interest requirement) (up to 5000 characters or 100 words):

If we develop a PRS for identifying people with hyperuricaemia at a high risk of gout, such people can be advised to modify their diet and lifestyle factors to reduce their gout risk. Likewise, if we develop a PRS for cardiovascular disease in gout, this will be used to educate patients with gout about their future risk of cardiovascular comorbidities and allow them to make dietary and lifestyle changes. This research aims to generate PRS models that have the potential

to identify new variants associated to gout, they could help elucidate pathways involved in evolution from hyperuricemia to gout.

A7. Please provide up to 6 keywords which best summarise your proposed research project:

Gout, hyperuricemia, polygenic risk score, genotyping

A8. Please provide a lay summary of your research project in plain English, stating the aims, scientific rationale, project duration and public health impact (up to 5000 characters or 400 words):

Gout affects one in forty people, and, manifests with episodes of severe joint pain and swelling that typically last for 1-2 weeks. If left untreated, the flares become more frequent, and, irreversible joint damage may result.

Gout occurs due to high urate levels in the body, which deposit inside the joints as urate crystals. Gout flares occur when these crystals are shed. However, only one in ten people with high urate level develops gout. The reasons why only a minority of people with high urate levels develop gout are not well understood. Moreover, knowing which person with high urate level has a high risk of gout will help prevent gout in those at high risk. Apart from this, research suggests that gout associates with high blood pressure, and heart disease. However, it is unknown which individual with gout is at risk of these comorbidities.

Thus, the purpose of this study is to develop a risk score that can identify people with gout from those without gout, and those without gout but with high blood urate levels. Finally, we will also develop a risk score to identify people with gout at risk of developing high blood pressure and heart diseases.

In this study, we will calculate a genetic risk score using all the information from common and rare genetic variants in participants of the UK biobank. Next we will add in information about other lifestyle factors, comorbidities, and medications to find out if this improves accuracy of the prediction score.

In order to generate these prediction scores, we will use data from people with disease of interest i.e. gout or gout and heart disease respectively and controls without these conditions. The controls will be matched to cases for age and sex.

This research will be completed in 2 years and will form part of the PhD thesis of Ms. Sandoval-Plata. The findings of this study will allow people with high

urate levels to find out their risk of developing gout, and will help people with gout understand their risk of developing heart disease and hypertension.
People with high risk can be advised to make lifestyle and dietary changes to prevent these illnesses.
This fits the UK Biobank's purpose of improving the prevention, diagnosis and treatment of life-threatening diseases.

A9. Will the research project result in the generation of any new data fields derived from existing complex datasets, such as imaging, accelerometry, electrocardiographic, linked healthcare data, etc, which might be of significant utility to other researchers:

☐ Yes

☑ No

A10. What is the estimated duration of your project, in months? If you consider (because for example the project is one involving the generation of hypotheses) that it would be difficult to set a fixed end point, we are prepared to consider a rolling 3-year period (during which annual updates are required):
24 months

Please note that you are expected to publish (or to make publicly available) your results and return to UK Biobank:

- any important derived variables
- a description of the methods used to generate them
- the underlying syntax/code used to generate the main results of the paper, and
- a short layman's description that summarises your findings.

These should be provided within six months of each publication or within 12 months of the project end date (whichever comes first). We also ask that you send us a copy of your accepted manuscript at least two weeks prior to publication and alert us if there are any ethical or contentious issues surrounding the findings.