Sumpter Nicholas Andrew (Orcid ID: 0000-0002-0451-9024)
Kapetanovic Meliha C (Orcid ID: 0000-0002-7853-514X)
Matsuo Hirotaka (Orcid ID: 0000-0001-5021-1613)
Roddy Edward (Orcid ID: 0000-0002-8954-7082)
Stiburkova Blanka (Orcid ID: 0000-0003-2465-1476)
Uhlig Till (Orcid ID: 0000-0002-6881-9552)
Stamp Lisa Katrina (Orcid ID: 0000-0003-0138-2912)
Dalbeth Nicola (Orcid ID: 0000-0003-4632-4476)
Merriman Tony R (Orcid ID: 0000-0003-0844-8726)

Association of Gout Polygenic Risk Score with Age at Disease Onset and Tophaceous Disease in European and Polynesian Men with Gout

Nicholas A Sumpter, MSc[1, 2], Riku Takei, MSc[1], Murray Cadzow, PhD[2], Ruth K G Topless, BSc[2], Amanda J Phipps-Green, MSc[2], Rinki Murphy, MBChB, PhD[3], Janak de Zoysa, MBChB, MClinEd[3], Huti Watson, MPH[4], Muhammad Qasim, PhD[4], Alexa S Lupi, MS[5, 6], Abhishek Abhishek, MD, PhD[7, 8], Mariano Andrés, MD, PhD[8, 9], Tania O Crișan, MD, PhD[8, 10], Michael Doherty, MD[7, 8], Lennart Jacobsson, MD, PhD[8, 11], Matthijs Janssen, MD, PhD[8, 12], Tim L Jansen, MD, PhD[8, 12], Leo A B Joosten, PhD[8, 10, 13], Meliha Kapetanovic, MD, PhD[8, 14], Frédéric Lioté, MD, PhD[8, 15], Hirotaka Matsuo, MD, PhD[8, 16], Geraldine M McCarthy, MD[8, 17], Fernando Perez-Ruiz, MD, PhD[8, 18], Philip Riches, MBChB, PhD[8, 19], Pascal Richette, MD, PhD[8, 15], Edward Roddy, DM[8, 20], Blanka Stiburkova, PhD[8, 21, 22], Alexander So, PhD, FRCP[8, 23], Anne-Kathrin Tausche, MD[8, 24], Rosa J Torres, MD, PhD[8, 25], Till Uhlig, MD, PhD[8, 26], Tanya J Major, PhD[2], Lisa K Stamp, MBChB, PhD[8, 27], Nicola Dalbeth, MD[3, 8], Hyon K Choi, MD, PhD[28], Ana I Vazquez, PhD[5, 6], Megan P Leask, PhD[1], Richard J Reynolds, PhD[1*], Tony R Merriman, PhD[1, 2, 8*^]

1. Division of Clinical Immunology and Rheumatology, Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA
2. Department of Biochemistry, University of Otago, Dunedin, New Zealand

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/art.42393

3. Department of Medicine, Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand
4. Ngāti Porou Hauora Trust, Te Puia Springs, New Zealand
5. Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, Michigan, USA
6. The Institute for Quantitative Health Science & Engineering, Michigan State University, East Lansing, Michigan, USA
7. Academic Rheumatology, School of Medicine, University of Nottingham, Nottingham City Hospital, Nottingham, UK
8. GlobalGoutGenetics Consortium
9. Department of Rheumatology, Alicante General University Hospital-ISABIAL, Miguel Hernandez University, Alicante, Spain
10. Department of Medical Genetics, Iuliu Haţieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania
11. Department of Rheumatology & Inflammation Research, Institute of Medicine, The Sahlgrenska Academy, University of Gothenburg, S-41346 Sweden
12. Complex Gout Expert Centre, Department of Rheumatology, Viecuri Medical Centre, Venlo, The Netherlands
13. Department of Internal Medicine, Radboud Institute of Molecular Life Sciences (RIMLS), Radboud University Medical Center, Nijmegen, The Netherlands
14. Lund University and Skåne University Hospital, Lund, Sweden
15. Hôpital Lariboisière, Assistance Publique-Hopitaux de Paris, and INSERM UMR-1132 and Université Paris Cité, Paris, France

16. Department of Integrative Physiology and Bio-Nano Medicine, National Defense Medical College, Saitama 359-8513, Japan
17. Mater Misericordiae University Hospital and University College Dublin, Ireland
18. Osakidetza, OSI EE-Cruces, Cruces University Hospital, Rheumatology Division, Biocruces-Bizkaia Health Research Institute, and Medicine Department of the Medicine School of the University of the Basque Country
19. IGC, University of Edinburgh, Scotland
20. School of Medicine, Keele University, Keele, Staffordshire, ST5 5BG, UK
21. Department of Pediatrics and Inherited Metabolic Disorders, First Faculty of Medicine, Charles University and General University Hospital, Prague, Czech Republic
22. Institute of Rheumatology, Prague, Czech Republic
23. Service of Rheumatology, Department of Musculoskeletal Medicine, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne, Switzerland
24. Division of Rheumatology, University Clinic Carl Gustav Carus at the TU Dresden, Dresden, Germany
25. Department of Biochemistry, La Paz University Hospital Health Research Institute (FIBHULP), IdiPaz, and Center for Biomedical Network Research on Rare Diseases (CIBERER), ISCIII, Madrid, Spain.
26. Center for Treatment of Rheumatic and Musculoskeletal Diseases (REMEDY), Diakonhjemmet Hospital, Oslo, Norway
27. Department of Medicine, University of Otago, Christchurch, Christchurch, New Zealand
28. Clinical Epidemiology Unit, Massachusetts General Hospital, Boston, Massachusetts, USA

* These authors contributed equally

^ Corresponding author (Email: tony.merriman@otago.ac.nz; Street Address: Room 202, 1825 University Blvd, Birmingham, AL, USA, 35233)

Funding: This work was supported by the University of Alabama at Birmingham, the University of Otago, and the Health Research Council of New Zealand.

Abstract:

Objective

To determine whether a gout polygenic risk score (PRS) associates with age at gout onset and tophaceous disease in European (EUR), East Polynesian (EP), and West Polynesian (WP) men/women with gout.

Methods

A 19-variant gout PRS was produced in seven European gout cohorts (N_{TOTAL} = 4,016), two East Polynesian gout cohorts (N_{TOTAL} = 682), and one West Polynesian gout cohort (N = 490). Sex-stratified regression models were used to estimate the relationship between the PRS and age at onset/tophaceous disease.

Results

The PRS associated with earlier age at gout onset in men (β_{EUR} [95%-CI]: -3.61 [-4.32, -2.90] years per unit PRS; β_{EP}: -6.35 [-8.91, -3.80]; β_{WP}: -3.51 [-5.46, -1.57]) but not in women (β_{EUR}: 0.07 [-2.32, 2.45]; β_{EP}: 0.20 [-7.21, 7.62]; β_{WP}: -3.33 [-9.28, 2.62]). The PRS showed a positive association with tophaceous disease in men (OR_{EUR}: 1.15 [1.00, 1.31]; OR_{EP}: 2.60 [1.66, 4.06]; OR_{WP}: 1.53 [1.07, 2.19]) but not in women (OR_{EUR}: 0.68 [0.42, 1.10]; OR_{EP}: 1.45 [0.39, 5.36]). The age at onset association was robust to the removal of *ABCG2* variants from the PRS in Europeans and East Polynesians but not West Polynesians (β_{EUR}: -2.42 [-3.37, -1.46]; β_{EP}: -6.80 [-10.06, -3.55]; β_{WP}: -1.79 [-4.74, 1.16]).

Conclusion

Gout risk variants also harbor risk for earlier age at onset and tophaceous disease in European and Polynesian men. Our findings suggest that earlier gout onset involves the

accumulation of gout risk alleles in men but perhaps not women, and that this genetic risk is shared across multiple ancestral groups.

Introduction

Gout is characterized by an inflammatory response to monosodium urate (MSU) crystal deposition that occurs in hyperuricemia. There is heterogeneity within gout presentations, with some people experiencing more severe disease as reflected by earlier age at onset and development of tophaceous disease [1, 2]. The source of this heterogeneity is unclear, although ethnicity and sex appear to be involved [3]. There is also evidence that genetic variants at the *ABCG2* locus have a role in age at onset and tophaceous disease in multiple ancestral groups [4-6]. Further clarification of the causes of variation in gout severity is important for management, as gout severity associates with lower quality of life and increased mortality [7, 8]. Throughout this study, we use the phrase "severity traits" to refer to both age at onset and tophaceous disease.

Existing genetic studies of gout severity have primarily focused on a few of the most strongly associated serum urate variants. While there is a clear genetic influence on serum urate concentrations [9], crystallization and inflammation likely also have a genetic underpinning [10, 11]. For this reason, it may be important to identify genetic variants from gout GWAS rather than serum urate GWAS, in case these variants have an effect on gout independent of their effect on serum urate [11]. Furthermore, previous studies have been limited to only testing the effects of individual variants, which likely leads to a lack of statistical power when these variants have small effect sizes [12]. The combination of individual variants into polygenic risk scores (PRS) has become increasingly common and has potential for direct use as a prediction tool [13], while also increasing statistical power for detecting genetic associations. There is also a substantial need for more sex-specific and multi-ancestral studies, such that research efforts can benefit a wider demographic and can potentially become useful for personalized medicine [14]. Therefore,

the aim of this study was to establish whether gout-associated genetic variants combined into a polygenic risk score associate with younger age at gout onset and tophaceous disease in sex- and ancestry-stratified analyses.

Materials and Methods

Cohort information

To establish a gout PRS, a total of 332,370 individuals of European ancestry from the UK Biobank were included in a gout GWAS. This cohort was recruited from the general population of the United Kingdom, with detailed phenotypic information available [15]. Detailed information on gout ascertainment and inclusion/exclusion criteria can be found in Supplementary Methods 1 & 2. In total, 7,131 gout cases and 325,239 controls were included for this analysis. This cohort was genotyped using the Affymetrix UK Biobank Axiom® array, followed by genotype imputation with IMPUTE4 using both the HRC reference panel and the 1000G + UK10K reference panel [15].

Seven European gout cohorts were included in the genetic analysis of the severity of gout (Table 1, S2 & S3): European individuals from the Genetics of Gout in Aotearoa study and five smaller studies in Australia and Aotearoa New Zealand (the Aus/NZ European cohort) [4]; the GlobalGout (formerly EuroGout) cohort [4]; and five Ardea Biosciences (Ardea) cohorts, sourced from urate-lowering therapy clinical trials funded and run by Ardea Biosciences Inc.: LASSO [16], CLEAR 1 [17], CLEAR 2 [18], CRYSTAL [19], and LIGHT [20]. The Aus/NZ European cohort also included non-gout controls.

Three Polynesian cohorts were also included in this study, largely sourced from the Genetics of Gout in Aotearoa study [21, 22] (Table 1, S2, & S3). These three cohorts included one West Polynesian cohort and two East Polynesian cohorts. The two East Polynesian cohorts consisted of those recruited in collaboration with the Ngāti Porou Hauora (Health Service) Charitable Trust and those recruited elsewhere in Aotearoa New Zealand. The Ngāti Porou Hauora cohort was recruited from the rohe (area) of the Ngāti Porou iwi (tribe) of the Tairāwhiti

(East Coast of the North Island of New Zealand) region. East and West Polynesian cohorts were defined based on a combination of self-reported ethnicity and genetic principal component analysis of genome-wide genotype data [22], where East Polynesian refers to those of Aotearoa NZ Māori and Cook Island Māori ancestry (excluding the island of Pukapuka) and West Polynesian refers to those of Samoan, Tongan, Niuean, Tokelauan and Pukapukan ancestry. All three Polynesian cohorts also included non-gout controls.

Phenotypic information for all cohorts excluding the UK Biobank was obtained from a combination of self-report questionnaires, laboratory measurements, and in-person assessments. Specifically, age at onset was self-reported, and tophaceous disease was defined through a combination of self-report and physical examination for at least one subcutaneous tophus by a trained assessor. Any individuals missing sex, age, the PRS, or both severity phenotypes were excluded from the analysis (Table S2 & S3). Individuals that were related at the first degree or more (KING kinship coefficient > 0.177 [23]) were removed (including duplicate samples across studies). These cohorts were genotyped using the Illumina CoreExome 24 v1.0, v1.1, or v1.3 arrays, with a small subset being genotyped on the OmniExome 8 v1.3 array. Details on reliability of ascertainment of phenotypic information for all cohorts can be found in Supplementary Methods 1 & 2.

Genome-wide association study (GWAS) and polygenic risk score calculation

A gout GWAS was performed in the UK Biobank cohort using a total of 27,287,012 imputed or genotyped variants. A logistic regression model was produced for each variant, adjusting for age, sex and the first 40 genetic principal components using PLINK version 1.9 build 6.10 [24]. Variants were considered statistically significant if they had $P < 5 \times 10^{-8}$.

A total of 19 genetic variants at 15 loci were identified as conditionally independent SNPs associated with gout at genome-wide significance in the UK Biobank (Table S1; Supplementary Methods 1 & 2). Three of the 15 loci had more than one independent gout-associated SNP, these were *ABCG2* (2 SNPs), *SLC2A9/WDR1* (3 SNPs), and *SLC22A11/NRXN2* (2 SNPs). Genotype data for all 19 variants were used to produce a gout PRS in all cohorts. This PRS was calculated for each individual by summing the number of risk alleles multiplied by the respective gout effect size (log odds) for each of the 19 risk alleles. Importantly, these weightings were based on models conditioning on all other PRS SNPs at the same locus, ensuring the weights represented independent effects of SNPs. We produced a second PRS that excluded both variants at the *ABCG2* locus.

Given that there is uncertainty on the best way to measure gout diagnosis in epidemiological studies, e.g. ICD code vs physician diagnosis, but there is a standard clinical assay for serum urate, we decided to compare our gout PRS to a serum urate PRS. We produced a serum urate PRS that included 82 serum urate associated variants identified in the European GWAS of serum urate performed by Tin et al. (2019) that were also genotyped in our studied cohorts (Supplementary Methods 1 & 2; Table S6). We also performed a sensitivity analysis to determine the effect of the weighting method on associations with male age at onset (Supplementary Methods 1 & 2).

Statistical analysis

Logistic or linear regression models were used to test for association of the PRS with gout, age at gout onset, and tophaceous disease. All models were performed in a sex-stratified manner. The first 10 genetic principal components were used as covariates in all models, derived

from over 12,000 genotyped participants from the cohorts studied here (excluding the UK Biobank), including participants of other ancestral backgrounds, as previously described [22]. For all Polynesian models, an additional set of 10 genetic principal components were included as covariates. These were derived from a subset of the above 12,000 participants, only including those who self-reported being of any Oceanian ethnicity (Polynesian, Micronesian, or Melanesian). Gout models were run with and without adjustment for age at collection. Tophus models were run with and without adjustment for disease duration.

All European models were tested in a fixed-effect meta-analysis, with each model separately run and meta-analyzed across the seven cohorts using the "meta" package for R 4.0.2 [25, 26]. Fixed-effect meta-analyses were also performed over the two East Polynesian cohorts for age at onset and gout models, though the two cohorts were pooled for tophaceous disease models due to the Ngāti Porou cohort having too few tophus-positive cases to be modeled separately.

Given previously reported associations of variants at the *ABCG2* locus with both severity traits investigated here [4-6], we also regressed the two severity traits on both conditionally independent variants at the *ABCG2* locus, followed by models using the PRS with both *ABCG2* variants excluded. In all individual variant models, $P < 0.025$ (0.05/2 variants to account for multiple testing) was used as the threshold for establishing statistical significance, while $P < 0.05$ was used for establishing significance in all PRS models.

For age at onset models, cohorts with $N < 20$ (European women in the CLEAR 1, CLEAR 2, CRYSTAL, and LIGHT trial cohorts) were excluded as it was not possible to accurately include them in meta-analyses. For tophus models, cohorts with $N < 20$ for either tophus-positive cases or tophus-negative cases (European men from the CRYSTAL trial,

European women from all 5 Ardea cohorts, and West Polynesian women) were excluded, due to the inability for logistic models to converge.

To determine the relative explanatory power of the PRS in comparison to other risk factors for earlier age at gout onset, a model was produced with all gout cases, regressing age at onset on the PRS, sex, ancestry, and BMI, then calculating partial R^2 values for each predictor.

For comparison of the age at onset association for the standardized gout PRS vs a standardized urate PRS derived from Tin et al. (2019), we performed linear models in each pooled male gout cohort for each ancestry group. A PRS weighting sensitivity analysis was also performed in which the three differentially weighted 15-variant risk scores were each standardized in pooled male gout cohorts for each ancestral group, then modeled with age at onset as the outcome. Standardized effect sizes and partial R^2 values were compared for each PRS in both analyses.

Finally, to directly investigate any sex differences, we determined the standardized effect size of the gout PRS on age at onset within sex- and ancestry-stratified cohorts. We then repeated these models in ancestry-stratified cohorts along with an interaction term between the gout PRS and sex.

R-code used in all analyses can be found in Supplementary Methods 2 and on GitHub (https://github.com/MerrimanLab/PRS-Project).

Results

Descriptive statistics

Gout cohort sizes ranged from 117 to 1,188, with each cohort being 77% to 98% male (Tables 1 and S3). The Aus/NZ European and GlobalGout cohorts had similar mean age at recruitment, at 62 and 60 years for men respectively, and 70 and 68 years for women. Ardea cohorts consistently had mean age at recruitment between 51 and 54 years for men, and between 55 and 64 years for women. The two East Polynesian cohorts had mean age at recruitment of 54/60 years for men, and 61/59 years for women. The West Polynesian cohort had the youngest mean age at recruitment, with an average age of 48 years for men and 53 years for women. The male European gout cohorts had mean age at onset of gout between 40 and 47 years, while female European gout cohorts had mean disease onset between 46 and 62 years of age. The East and West Polynesian cohorts consistently had earlier mean age at onset, with men on average developing gout between 35 and 39 years, and women developing gout between 44 and 49 years. Mean disease duration ranged from 3 years for women in the Ardea CRYSTAL cohort to 22 years for men in the Ngāti Porou cohort. Most cohorts had between 10% and 60% of cases with at least one tophus present and negligible difference between sexes, though 100% of the CRYSTAL study cohort had tophaceous gout due to the study design. All Ardea cohorts and the male West Polynesian cohort had a median of between 3 and 5 self-reported flares in the last year, while East Polynesian, Aus/NZ European, GlobalGout, and female West Polynesian cohorts had median flare rates of between 1 and 3 in the last year. There were no obvious differences in self-reported flare frequency between sexes, except for an apparent increase in flares among males of the West Polynesian cohort. The mean PRS ranged from 4.0 to 4.3 for the European gout cohorts, 4.2 to 4.4 in the East Polynesian gout cohorts and 4.7 to 4.8 in the West

Polynesian gout cohort. Across all gout cohorts the PRS ranged from 1.7 to 6.5, with the bottom 10% of individuals below a value of 3.3 and the top 10% of individuals above a value of 4.9.

Gout polygenic risk score (PRS) derivation and validation

A total of 19 independent genetic variants were identified as associated with gout at genome-wide significance ($P < 5 \times 10^{-8}$; Figure S1; Table S1). All 19 variants were also associated with serum urate in people of European ancestry (Table S1) [9]. The PRS showed a positive association with gout in the Aus/NZ European case-control cohort, both before and after adjustment for age at collection (age-adjusted results: OR_{MALE} [95%-CI]: 2.66 [2.25, 3.15] per unit increase in PRS; OR_{FEMALE}: 1.94 [1.45, 2.60]) (Table S4). Additionally, the PRS associated with gout in East and West Polynesian case-control cohorts, both before and after adjustment for age at collection (age-adjusted results: East Polynesian: OR_{MALE}: 1.86 [1.31, 2.64]; OR_{FEMALE}: 1.91 [1.13, 3.23]; West Polynesian: OR_{MALE}: 3.84 [2.64, 5.57]; OR_{FEMALE}: 5.43 [2.63, 11.20]).

Association of PRS with age at onset and tophaceous disease

Meta-analysis of the European gout cohorts revealed a negative association between the PRS and age at onset in men (β: -3.61 [-4.32, -2.90] years per unit PRS) (Figure 1; Table S5). This negative association was also seen in both East and West Polynesian men (β_{EP}: -6.35 [-8.91, -3.80]; β_{WP}: -3.51 [-5.46, -1.57]) (Figure 1; Table S5). The equivalent models in women were all non-significant (β_{EUR}: 0.07 [-2.32, 2.45]; β_{EP}: 0.20 [-7.21, 7.62]; β_{WP}: -3.33 [-9.28, 2.62]) (Figure 1; Table S5). We found that the *ABCG2* variant rs2231142 was individually associated with age at onset in all three ancestral groups in men (β_{EUR}: -3.12 years per allele [-3.98, -2.27]; β_{EP}: -4.41 [-7.39, -1.44]; β_{WP}: -3.40 [-5.20, -1.61]) but not women (β_{EUR}: 1.24 [-1.68, 4.16]; β_{EP}: 3.43 [-

4.53, 11.38]; β_{WP}: 0.01 [-5.23, 5.26]) (Figure S2; Table S5). We also found that the *ABCG2* variant rs10011796 was associated with age at onset in European men independent of rs2231142 (β_{EUR}: -0.97 [-1.67, -0.27]; β_{EP}: 0.32 [-1.62, 2.26]; β_{WP}: 0.03 [-1.75, 1.81]) (Figure S2; Table S5). In a sensitivity analysis, removal of both *ABCG2* variants from the PRS resulted in a reduced effect size for the association of the PRS with age at onset in European men with gout, no change in East Polynesian men with gout, and a loss of association in West Polynesian men with gout (β_{EUR}: -2.42 [-3.37, -1.46]; β_{EP}: -6.80 [-10.06, -3.55]; β_{WP}: -1.79 [-4.74, 1.16]) (Figure 1; Table S5).

When age at onset was regressed on the PRS, sex, ancestry, and BMI within the pooled gout cohort, partial R^2 values showed that the PRS explained 2.9% of variance in age at onset, compared to 8.4%, 1.8%, and 3.0% for sex, ancestry, and BMI respectively.

The PRS showed a positive association with tophaceous disease in all male-stratified models (OR_{EUR}: 1.15 [1.00, 1.31]; OR_{EP}: 2.60 [1.66, 4.06]; OR_{WP}: 1.53 [1.07, 2.19]) (Figure 2; Table S5). Of these, only the East Polynesian model was robust to adjustment for disease duration (OR_{EUR}: 1.11 [0.97, 1.28]; OR_{EP}: 2.13 [1.32, 3.43]; OR_{WP}: 1.46 [0.99, 2.14]) (Figure 2; Table S5) or removal of *ABCG2* variants from the PRS (OR_{EUR}: 1.15 [0.96, 1.38]; OR_{EP}: 3.25 [1.87, 5.67]; OR_{WP}: 1.24 [0.73, 2.13]) (Figure S3; Table S5). The PRS did not associate with tophaceous disease in women (OR_{EUR}: 0.68 [0.42, 1.10]; OR_{EP}: 1.45 [0.39, 5.36]) (Figure 2; Table S5). Neither *ABCG2* variant was associated with tophaceous disease in any cohort (Figure S3; Table S5).

Comparing the gout PRS association with male age at onset to a serum urate PRS revealed a slightly reduced standardized effect in each ancestral group for the serum urate PRS, though the confidence intervals did overlap (European β_{GOUT}: -2.59 [-3.06, -2.13] vs. β_{SU}: -2.08

[-2.56, -1.60]; East Polynesian β_{GOUT}: -3.11 [-4.34, -1.88] vs. β_{SU}: -2.48 [-3.74, -1.22]; West Polynesian β_{GOUT}: -2.01 [-3.15, -0.87] vs. β_{SU}: -1.67 [-2.84, -0.51]). This was reflected in the higher partial R^2 values for the gout PRS (European: 3.2%, East Polynesian: 4.5%, West Polynesian: 2.6%) compared to the urate PRS (European: 2.1%, East Polynesian: 2.8%, West Polynesian: 1.7%).

In the PRS weighting sensitivity analysis, we found that all weighting schemes performed similarly for association with male age at onset in all three ancestral groups, though the gout weighted PRS models consistently showed the largest standardized effect sizes and partial R^2 values (European: β_{GOUT}: -2.48 [-2.94, -2.01], 3.0% partial R^2; β_{URATE}: -2.23 [-2.69, -1.76], 2.4% partial R^2; β_{UNW}: -1.69 [-2.16, -1.21], 1.4% partial R^2; East Polynesian: β_{GOUT}: -2.73 [-3.96, -1.50], 3.5% partial R^2; β_{URATE}: -2.55 [-3.78, -1.31], 3.0% partial R^2; β_{UNW}: -2.26 [-3.50, -1.03], 2.3% partial R^2; West Polynesian: β_{GOUT}: -1.93 [-3.06, -0.80], 2.4% partial R^2; β_{URATE}: -1.51 [-2.65, -0.38], 1.4% partial R^2; β_{UNW}: -1.48 [-2.62, -0.35], 1.3% partial R^2; Table S7).

Finally, our direct investigation of sex differences in age at onset identified a significant interaction between sex and the gout PRS for age at onset models in Europeans only (European $P_{INT} = 0.02$; East Polynesian $P_{INT} = 0.32$; West Polynesian $P_{INT} = 0.63$). Standardized effect sizes of the gout PRS in each sex showed some evidence for sex differences on age at onset in Europeans only (β_{MALE}: -2.58 [-3.04, -2.11]; β_{FEMALE}: -0.77 [-2.24, 0.70]). Neither East Polynesians (β_{MALE}: -3.07 [-4.28, -1.86]; β_{FEMALE}: -1.67 [-4.37, 1.03]) or West Polynesians (β_{MALE}: -2.01 [-3.15, -0.87]; β_{FEMALE}: -2.91 [-7.21, 1.40]) showed any difference between sexes.

Discussion

In this study, we have shown that a set of 19 European gout genetic risk variants combined into a PRS associates with earlier age at gout onset and tophaceous disease in men with gout from multiple ancestral backgrounds. We did not detect these associations in women with gout, which may suggest that age at gout onset in women is comparatively less determined by gout genetic risk factors. This is plausibly due to a protective role of lower premenopausal serum urate levels and thus other risk factors occurring in later life (i.e. diuretics, renal disease and/or type 2 diabetes) are more likely driving the risk for hyperuricemia and gout in women [27-32]. This sex difference is further evidenced by a significant interaction between sex and gout genetic risk for age at onset in Europeans. The PRS independently explained 2.9% of the variability in age at gout onset in all gout cases, which is comparable to the effect of BMI (as measured at recruitment) and ancestry in the same cohort, but less than half of the effect of sex on age at onset. In terms of clinical impact, the model results suggest that men with gout in the top 10% of gout genetic risk on average experience their first gout flare 5.8 to 10.6 years (2.6 to 14.9 years based on 95% confidence intervals) earlier than those in the bottom 10% of gout genetic risk, depending on their ancestral background. The rs2231142 variant at the *ABCG2* locus has previously been implicated in age at gout onset [4, 6, 33]. Here, we also show that this variant associates with age at onset, however we additionally show a novel independent effect of a second variant at this locus (rs10011796) on age at onset in male European gout cases. Further, we demonstrate that removal of both *ABCG2* variants from the PRS did not result in a loss of association for the PRS with age at onset in European and East Polynesian men. The PRS was shown to be associated with gout in both sexes within cohorts of European, East Polynesian, and West Polynesian ancestry. This study highlights the importance of stratifying studies by sex

when there are known sex differences in disease presentation and epidemiology. We also demonstrated the multi-ancestral context of these associations, highlighting a shared genetic cause of earlier onset gout that could have implications in other ancestral groups not included here.

Recently, Zaidi et al. (2020) showed an association between *ABCG2* rs2231142 and early onset gout in a subset of the cohorts in our study. In the present study, we have improved power to detect genetic associations with age at onset through the addition of over 1,000 more European gout cases and through combination of variants into a PRS rather than using individual variants in our models. By performing our analysis in a sex-stratified manner and testing for associations with age at onset as a linear variable rather than a binary variable, we were able to obtain more precise estimates of the effect of gout genetic risk factors on age at onset. Finally, we derived our genetic variants from a gout GWAS rather than a serum urate GWAS, enabling us to weight our PRS towards gout, which differs from weightings towards serum urate (Table S1; Table S7). Our PRS weighting sensitivity analyses showed that all identified associations were not dependent on the weighting scheme, though the gout-weighted risk scores appeared to have slightly better explanatory power than the serum urate-weighted and unweighted risk scores in all three ancestral groups. Further, a PRS derived from a serum urate GWAS, with 67 additional variants compared to the gout PRS, did not show evidence for an increased association with age at gout onset, highlighting the relevance of using a gout PRS instead of a serum urate PRS. A definitive answer as to whether a urate or gout PRS is more predictive of gout disease severity would require training models and then assessing their predictive performance in separate test datasets.

The PRS also associated with increased likelihood of tophaceous disease in men from all three ancestral backgrounds. Given the association of disease duration with tophaceous disease

[5], this association could be mediated by age at onset. By adjusting for disease duration in these models, we were able to show incomplete attenuation of this association in East Polynesian men, thus indicating a modest independent effect of the gout PRS on tophaceous disease in this cohort. A prior study of the effect of the *ABCG2* and *SLC2A9* loci on tophaceous gout in the Genetics of Gout in Aotearoa cohort showed evidence for a Polynesian-specific role of both *ABCG2* variants on this trait, specific to West Polynesians for the rs10011796 variant [5]. This is discordant with our results, however we used an additive model for testing the association of *ABCG2* variants with tophaceous disease, whereas He et al. (2017) used a dominance model. By reproducing our models using the same method, we were able to replicate the results of He et al. (2017) (Supplementary Methods 2), suggesting a West Polynesian specific dominant effect of both *ABCG2* variants on tophaceous disease.

A potential limitation of our case-only study design relates to index event (collider) bias [34]. This occurs when risk factors for incident disease are used in studies of disease progression/severity. This bias results in an induced association between risk factors (both measured and unmeasured) when observational studies are conditioned on a common effect or collider (e.g. gout). If the risk factor (PRS for gout in our case) causes disease incidence, but not progression (i.e. tophaceous disease) then there is a risk of an induced negative association due to collider bias [35]. If a risk factor (i.e. a PRS) causes both incident disease and disease progression then the collider effect would result in a bias of the estimated effect on progression towards the null. Here we expect that collider bias has resulted in an underestimate of the effect of the PRS on tophaceous disease. However, given that age at onset is a measure of disease onset (rather than progression), it is unlikely that collider bias has influenced the estimate of effect of the PRS on age at onset. Another potential limitation of our study is that we have relatively small

numbers of female gout cases compared to males, and if female age at gout onset and/or tophaceous disease are under genetic control by gout genetic risk factors, we may not be sufficiently powered to detect them. Finally, it is likely that other genetic variants contribute to gout severity, and this study only tested those known to associate with gout in Europeans of the UK Biobank cohort.

Flare frequency is another measure of disease severity that we had data available on, however we chose not to examine the association of the PRS with this trait. One reason for this is that flare data was not consistently collected in each cohort, with all Ardea cohorts being selected for those with ≥ 2 flares in the last year. Additionally, many other variables contribute to flare frequency, including use of anti-inflammatory medications and urate lowering therapy, and ideally the trait should be measured with a combination of multiple criteria [36, 37]. Thus, if flare frequency indeed has a genetic basis, it will require careful measurement and a well-controlled study to investigate.

Our results may relate to the concept of primary vs secondary gout [38]. Generally, primary disease presents independent of other diseases, while secondary disease is acquired as a result of other diseases, such as via metabolic changes induced by obesity (i.e. insulin resistance). Primary disease will typically present at an earlier age than secondary disease, as secondary disease requires development of another disease first. Given this expected correlation of primary gout with earlier age at onset, our results may reflect an increased genetic risk for gout in men with primary gout versus those with secondary gout. Additionally, this higher genetic risk in primary gout and corresponding earlier age at onset would result in longer disease duration and higher likelihood of developing tophaceous disease (which appears to also be directly affected by gout genetic risk factors). Our results suggest that this is a male-specific

phenomenon, however our PRS is likely biased towards male gout genetic risk given the much higher prevalence of men with gout in the UK Biobank gout GWAS presented here. Therefore, if this also occurred in women, but due to different genetic risk factors, we would not expect to capture this effect. Also, as mentioned previously, we may be underpowered to detect associations in women with gout.

In conclusion, genetic risk for gout derived from a cohort of European ancestry associates with earlier age at gout onset and tophaceous disease in men from three ancestral backgrounds. We hypothesize that this reflects a trans-ancestral male-specific genetic predisposition to gout among primary gout cases that increases risk of tophaceous gout.

Conflicts

The authors declare no conflicts of interest.

Acknowledgements

This research has been conducted using the UK Biobank Resource (project ID: 12611). This research received funding from the New Zealand Health Research Council and Lottery Health. We would like to thank Ria Akuhata-Brown, Nancy Aupouri, Jill Drake, Carol Ford, Roddi Laurence, Meaghan House and Gabrielle Sexton for assistance in recruitment for the Genetics of Gout in Aotearoa study. Additionally, we would also like to acknowledge the European Crystal Network, formed after the first European Crystal Workshop in Paris, March 2010 (Prof Frédéric Lioté, Paris, and Prof Alexander So, Lausanne, convenors).

Figure Legends

Figure 1: Associations of a gout polygenic risk score with age at gout onset, either including *ABCG2* SNPs (A) or excluding *ABCG2* SNPs (B). Models were run within each unique combination of sex and ancestry. A fixed-effect meta-analysis was performed across all European cohorts within each sex, and separately in East Polynesian cohorts within each sex. NP = Ngāti Porou.

Figure 2: Associations of a gout polygenic risk score with tophaceous disease, either not adjusted for disease duration (A), or adjusted for disease duration (B). Models were run within each unique combination of sex and ancestry. A fixed-effect meta-analysis was performed across all European cohorts within each sex. Polynesian models were not meta-analyzed. The Ngāti Porou cohort was pooled with other East Polynesians for this analysis (see Methods).

Tables

Table 1: Relevant demographic and clinical characteristics of each gout cohort, separated by sex. Cohorts with fewer than 20 individuals not shown. Full summaries of all cohorts (including missing data percentages) can be found in Table S2 and S3.

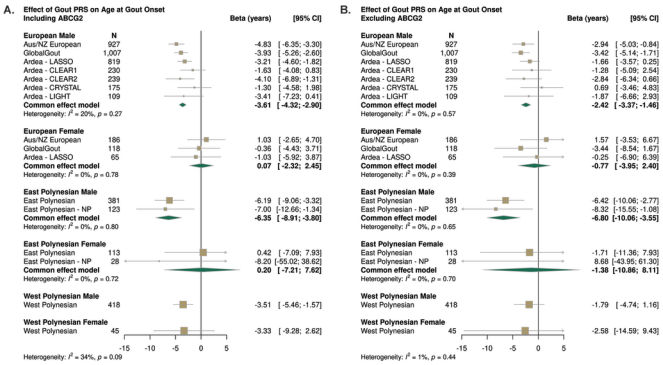
Cohort	Aus/NZ European		GlobalGout		Ardea LASSO		Ardea CLEAR 1	Ardea CLEAR 2	Ardea CRYSTAL	Ardea LIGHT	East Polynesian Ngāti Porou Hauora		West Polynesian	
Sex	Male	Female	Male	Female	Male	Female	Male	Male	Male	Male	Male	Female	Male	Female
N	978	210	1,032	124	819	65	230	239	175	109	408 124	122 28	436	54
Age at recruitment, years	62.4 ± 12.4	70.0 ± 12.7	60.1 ± 13.1	67.6 ± 11.1	51.4 ± 11.8	60.7 ± 10.6	52.3 ± 11.1	53.0 ± 10.8	53.9 ± 11.0	53.3 ± 11.8	54.3 ± 12.4 59.7 ± 11.3	60.7 ± 11.7 59.1 ± 13.3	47.5 ± 12.3	53.4 ± 13.4
Age at onset, years	46.4 ± 15.8	59.5 ± 15.7	46.5 ± 14.0	57.8 ± 12.5	41.4 ± 13.4	55.1 ± 12.0	41.9 ± 12.4	42.6 ± 13.2	40.1 ± 13.0	42.4 ± 13.1	37.9 ± 14.0 39.1 ± 15.2	49.4 ± 15.4 46.0 ± 16.8	34.6 ± 12.0	44.3 ± 15.0
Disease duration,	16.8 ± 12.7	10.9 ± 10.4	14.5 ± 11.4	10.6 ± 9.8	11.0 ± 9.4	6.6 ± 8.0	11.4 ± 9.4	11.4 ± 9.8	14.8 ± 10.0	11.9 ± 8.7	17.2 ± 12.8 21.7 ± 15.3	13.1 ± 13.2 14.1 ± 12.6	13.6 ± 10.3	9.2 ± 9.2
Flares in previous year	2 (0 - 4)	1.5 (0 - 3.25)	2 (1 - 4)	2.5 (1 - 4)	4 (3 - 8)	3 (3 - 6)	3 (2 - 6)	4 (2 - 8)	4 (3 - 6)	4 (2 - 10)	3 (1 - 6) 2 (0 - 3)	2 (0 - 5) 3 (1 - 6)	4 (2 - 10)	2 (1 - 5)
Tophaceous disease	333 (43.4)	67 (39.9)	320 (57.6)	46 (62.2)	138 (16.8)	5 (7.7)	34 (14.9)	54 (22.6)	174 (99.4)	26 (23.9)	144 (41.3) 9 (12.2)	26 (28.3) 4 (19.0)	177 (44.6)	14 (28.6)
Gout PRS	4.1 ± 0.7	4.0 ± 0.6	4.0 ± 0.6	4.0 ± 0.6	4.1 ± 0.7	4.1 ± 0.6	4.2 ± 0.7	4.2 ± 0.6	4.2 ± 0.6	4.1 ± 0.6	4.4 ± 0.5 4.2 ± 0.5	4.4 ± 0.5 4.4 ± 0.5	4.8 ± 0.6	4.7 ± 0.6

References

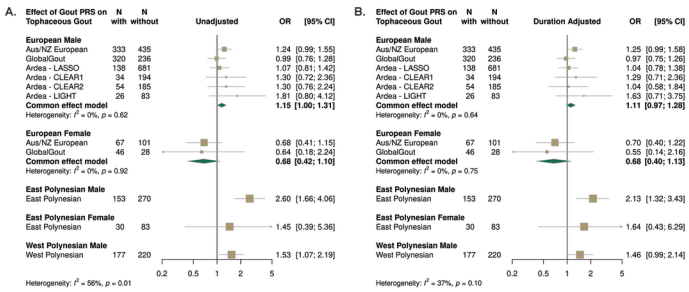
1. Zhang B, Fang W, Zeng X, Zhang Y, Ma Y, Sheng F, Zhang X: **Clinical characteristics of early- and late-onset gout: A cross-sectional observational study from a Chinese gout clinic**. *Medicine (Baltimore)* 2016, **95**(47):e5425.
2. Pascart T, Norberciak L, Ea HK, Guggenbuhl P, Liote F: **Patients With Early-Onset Gout and Development of Earlier Severe Joint Involvement and Metabolic Comorbid Conditions: Results From a Cross-Sectional Epidemiologic Survey**. *Arthritis Care Res (Hoboken)* 2019, **71**(7):986-992.
3. Kuo C-F, Grainge MJ, Zhang W, Doherty M: **Global epidemiology of gout: prevalence, incidence and risk factors**. *Nat Rev Rheumatol* 2015, **11**(11):649-.
4. Zaidi F, Narang RK, Phipps-Green A, Gamble GG, Tausche A-K, So A, Riches P, Andres M, Perez-Ruiz F, Doherty M *et al*: **Systematic genetic analysis of early-onset gout: ABCG2 is the only associated locus**. *Rheumatology* 2020:kez685.
5. He W, Phipps-Green A, Stamp LK, Merriman TR, Dalbeth N: **Population-specific association between ABCG2 variants and tophaceous disease in people with gout**. *Arthritis Res Ther* 2017, **19**(1):43.
6. Matsuo H, Ichida K, Takada T, Nakayama A, Nakashima H, Nakamura T, Kawamura Y, Takada Y, Yamamoto K, Inoue H *et al*: **Common dysfunctional variants in ABCG2 are a major cause of early-onset gout**. *Sci Rep* 2013, **3**(1):2014.
7. Hirsch JD, Terkeltaub R, Khanna D, Singh J, Sarkin A, Shieh M, Kavanaugh A, Lee SJ: **Gout disease-specific quality of life and the association with gout characteristics**. *Patient Relat Outcome Meas* 2010, **2010**:1-8.
8. Perez-Ruiz F, Martinez-Indart L, Carmona L, Herrero-Beites AM, Pijoan JI, Krishnan E: **Tophaceous gout and high level of hyperuricaemia are both associated with increased risk of mortality in patients with gout**. *Ann Rheum Dis* 2014, **73**(1):177-182.
9. Tin A, German Chronic Kidney Disease S, Lifelines Cohort S, Program VAMV, Marten J, Halperin Kuhns VL, Li Y, Wuttke M, Kirsten H, Sieber KB *et al*: **Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels**. *Nat Genet* 2019, **51**(10):1459-1474.
10. Bodofsky S, Merriman TR, Thomas TJ, Schlesinger N: **Advances in our understanding of gout as an auto-inflammatory disease**. *Semin Arthritis Rheum* 2020, **50**(5):1089-1100.
11. Major TJ, Dalbeth N, Stahl EA, Merriman TR: **An update on the genetics of hyperuricaemia and gout**. *Nat Rev Rheumatol* 2018, **14**(6):341-353.
12. Sham PC, Purcell SM: **Statistical power and significance testing in large-scale genetic studies**. *Nat Rev Genet* 2014, **15**(5):335-346.
13. Elliott J, Bodinier B, Bond TA, Chadeau-Hyam M, Evangelou E, Moons KGM, Dehghan A, Muller DC, Elliott P, Tzoulaki I: **Predictive Accuracy of a Polygenic Risk Score-Enhanced Prediction Model vs a Clinical Risk Score for Coronary Artery Disease**. *JAMA* 2020, **323**(7):636-645.
14. Peterson RE, Kuchenbaecker K, Walters RK, Chen CY, Popejoy AB, Periyasamy S, Lam M, Iyegbe C, Strawbridge RJ, Brick L *et al*: **Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations**. *Cell* 2019, **179**(3):589-603.
15. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J *et al*: **The UK Biobank resource with deep phenotyping and genomic data**. *Nature* 2018, **562**(7726):203-209.
16. Becker MA, Fitz-Patrick D, Choi HK, Dalbeth N, Storgard C, Cravets M, Baumgartner S: **An open-label, 6-month study of allopurinol safety in gout: The LASSO study**. *Semin Arthritis Rheum* 2015, **45**(2):174-183.

17. Saag KG, Fitz-Patrick D, Kopicko J, Fung M, Bhakta N, Adler S, Storgard C, Baumgartner S, Becker MA: **Lesinurad Combined With Allopurinol: A Randomized, Double-Blind, Placebo-Controlled Study in Gout Patients With an Inadequate Response to Standard-of-Care Allopurinol (a US-Based Study)**. *Arthritis & Rheumatology* 2017, **69**(1):203-212.
18. Bardin T, Keenan RT, Khanna PP, Kopicko J, Fung M, Bhakta N, Adler S, Storgard C, Baumgartner S, So A: **Lesinurad in combination with allopurinol: a randomised, double-blind, placebo-controlled study in patients with gout with inadequate response to standard of care (the multinational CLEAR 2 study)**. *Ann Rheum Dis* 2017, **76**(5):811-820.
19. Dalbeth N, Jones G, Terkeltaub R, Khanna D, Kopicko J, Bhakta N, Adler S, Fung M, Storgard C, Baumgartner S *et al*: **Lesinurad, a Selective Uric Acid Reabsorption Inhibitor, in Combination With Febuxostat in Patients With Tophaceous Gout: Findings of a Phase III Clinical Trial**. *Arthritis & Rheumatology* 2017, **69**(9):1903-1913.
20. Tausche A-K, Alten R, Dalbeth N, Kopicko J, Fung M, Adler S, Bhakta N, Storgard C, Baumgartner S, Saag K: **Lesinurad monotherapy in gout patients intolerant to a xanthine oxidase inhibitor: a 6 month phase 3 clinical trial and extension study**. *Rheumatology* 2017, **56**(12):2170-2178.
21. Ji A, Shaukat A, Takei R, Bixley M, Cadzow M, Topless RK, Major TJ, Phipps-Green A, Merriman ME, Hindmarsh JH *et al*: **Aotearoa New Zealand Maori and Pacific Population-amplified Gout Risk Variants: CLNK Is a Separate Risk Gene at the SLC2A9 Locus**. *J Rheumatol* 2021, **48**(11):1736-1744.
22. Krishnan M, Major TJ, Topless RK, Dewes O, Yu L, Thompson JMD, McCowan L, de Zoysa J, Stamp LK, Dalbeth N *et al*: **Discordant association of the CREBRF rs373863828 A allele with increased BMI and protection from type 2 diabetes in Maori and Pacific (Polynesian) people living in Aotearoa/New Zealand**. *Diabetologia* 2018, **61**(7):1603-1613.
23. **KING Tutorial: Relationship Inference** [https://www.kingrelatedness.com/manual.shtml]
24. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ *et al*: **PLINK: a tool set for whole-genome association and population-based linkage analyses**. *Am J Hum Genet* 2007, **81**(3):559-575.
25. R Core Team: **R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.** 2022.
26. Balduzzi S, Rucker G, Schwarzer G: **How to perform a meta-analysis with R: a practical tutorial**. *Evid Based Ment Health* 2019, **22**(4):153-160.
27. Richette P, Clerson P, Périssin L, Flipo R-M, Bardin T: **Revisiting comorbidities in gout: a cluster analysis**. *Ann Rheum Dis* 2015, **74**(1):142-147.
28. Te Kampe R, Janssen M, van Durme C, Jansen TL, Boonen A: **Sex Differences in the Clinical Profile Among Patients With Gout: Cross-sectional Analyses of an Observational Study**. *J Rheumatol* 2021, **48**(2):286-292.
29. Hak AE, Curhan GC, Grodstein F, Choi HK: **Menopause, postmenopausal hormone use and risk of incident gout**. *Ann Rheum Dis* 2010, **69**(7):1305-1309.
30. Sumpter N, Cadzow M, So A, Reynolds R, Merriman T: **Analysis of common gout comorbidities in the UK Biobank cohort reveals sex-specific effects and genetic differentiation**. In: *Arthritis & Rheumatology: 2020*: WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA; 2020.
31. Singh JA, Gaffo A: **Gout epidemiology and comorbidities**. *Seminars in Arthritis and Rheumatism* 2020, **50**(3, Supplement):S11-S16.
32. Patel AV, Gaffo AL: **Managing Gout in Women: Current Perspectives**. *J Inflamm Res* 2022, **15**:1591-1598.
33. Stiburkova B, Pavelcova K, Pavlikova M, Jesina P, Pavelka K: **The impact of dysfunctional variants of ABCG2 on hyperuricemia and gout in pediatric-onset patients**. *Arthritis Res Ther* 2019, **21**(1):77.

34. Choi HK, Nguyen US, Niu J, Danaei G, Zhang Y: **Selection bias in rheumatic disease research**. *Nat Rev Rheumatol* 2014, **10**(7):403-412.
35. Paternoster L, Tilling K, Davey Smith G: **Genetic epidemiology and Mendelian randomization for informing disease therapeutics: Conceptual and methodological challenges**. *PLoS Genet* 2017, **13**(10):e1006944.
36. Gaffo AL, Schumacher HR, Saag KG, Taylor WJ, Dinnella J, Outman R, Chen L, Dalbeth N, Sivera F, Vazquez-Mellado J *et al*: **Developing a provisional definition of flare in patients with established gout**. *Arthritis Rheum* 2012, **64**(5):1508-1517.
37. Gaffo AL, Dalbeth N, Saag KG, Singh JA, Rahn EJ, Mudano AS, Chen YH, Lin CT, Bourke S, Louthrenoo W *et al*: **Brief Report: Validation of a Definition of Flare in Patients With Established Gout**. *Arthritis Rheumatol* 2018, **70**(3):462-467.
38. Wallace SL, Robinson H, Masi AT, Decker JL, McCarty DJ, Yu TF: **Preliminary criteria for the classification of the acute arthritis of primary gout**. *Arthritis Rheum* 1977, **20**(3):895-900.



ART_42393_1.tiff



ART_42393_2.tiff