HHS Public Access

Author manuscript

Kidney Int. Author manuscript; available in PMC 2021 November 01.

Published in final edited form as:

Kidney Int. 2020 November ; 98(5): 1323–1330. doi:10.1016/j.kint.2020.04.055.

Derivation and Validation of Genome Wide Polygenic Score for Urinary Tract Stone Diagnosis

Ishan Paranjpe[1,2,3,4], **Noah Tsao**[5], **Renae Judy**[5], **Manish Paranjpe**[6], **Kumardeep Chaudhary**[1,3,4], **Derek Klein**[7,8,9], **Iain Forrest**[1,14], **Ross O'Hagan**[1,3,4], **Arjun Kapoor**[2], **John Pfail**[2], **Suraj Jaladanki**[1,3,4], **Fayzan Chaudhry**[4], **Akhil Vaid**[1,4], **Phan Q. Duy**[10], **CBIPM Genomics Team**[1], **Regeneron Genomics Team**[11], **John Cijiang He**[12,13], **Benjamin S Glicksberg**[4,14], **Steven G Coca**[12], **Mantu Gupta**[2], **Ron Do**[1,3,14], **Scott M. Damrauer**[5], **Girish N. Nadkarni**[1,3,4,12,13]

[1.]The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[2.]Department of Urology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[3.]The Bio*Me* Phenomics Center, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[4.]The Hasso Plattner Institute for Digital Health at Mount Sinai, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[5.]Department of Surgery, Perelman School of Medicine, University of Pennsylvania, Philadelphia.

[6.]Harvard Medical School, Boston, MA, USA

[7.]Malcolm Randall VA Medical Center, Gainesville, FL

[8.]Division of Vascular Surgery and Endovascular Therapy, University of Florida College of Medicine, Gainesville, FL

[9.]Broad Institute of Harvard and MIT

[10.]Medical Scientist Training Program, Yale University School of Medicine, New Haven, CT, USA.

[11.]Regeneron Genetics Center, Tarrytown, NY, USA.

[12.]Division of Nephrology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[13.]Renal Program, James J Peters VA Medical Center at Bronx, New York, NY, USA

Corresponding Author: Girish Nadkarni, MD, MPH, One Gustave Levy Place, Box 1243, New York, NY-10029, Phone Number: 212-241-1385, girish.nadkarni@mountsinai.org.

Author Contributions:

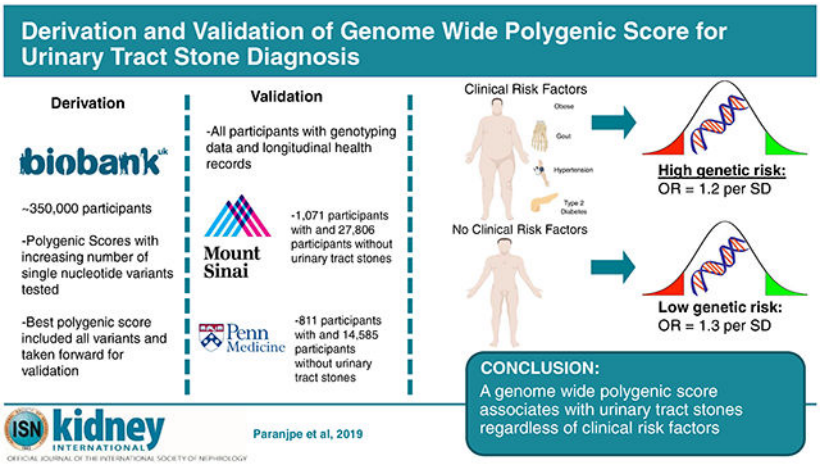
IP, GN, LC, JCH, SGC, RD, BG, MG, and SMD designed the study. IP, NT, RJ, GN, and RD acquired the data. IP, MP, RO, JP, AK, MG, KC, AV, FC, DK, NT, and RJ analyzed and interpreted the data. IP, PQ, SJ, MP, MG, GN, SGC, RD, LC, SMD, and JCH drafted and revised the manuscript.

[14]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Abstract

Urinary tract stones have high heritability indicating a strong genetic component. However, genome wide association studies (GWAS) have uncovered only a few genome wide significant single nucleotide polymorphisms (SNPs). Polygenic risk scores (PRS) sum cumulative effect of many SNPs and shed light on underlying genetic architecture. Using GWAS summary statistics from 361,141 participants in the United Kingdom Biobank, we generated a PRS and determined association with stone diagnosis in 28,877 participants in the Mount Sinai Bio*Me* Biobank. In Bio*Me* (1,071 cases and 27,806 controls), for every standard deviation increase, we observed a significant increment in adjusted odds ratio of a factor of 1.2 (95% confidence interval 1.13–1.26). In comparison, a risk score comprised of GWAS significant SNPs was not significantly associated with diagnosis. After stratifying individuals into low and high-risk categories on clinical risk factors, there was a significant increment in adjusted odds ratio of 1.3 (1.12–1.6) in the low- and 1.2 (1.1–1.2) in the high-risk group for every standard deviation increment in PRS. In a 14,348-participant validation cohort (Penn Medicine Biobank), every standard deviation increment was associated with a significant adjusted odds ratio of 1.1 (1.03 – 1.2). Thus, a genome wide PRS is associated with urinary tract stones overall and in the absence of known clinical risk factors and illustrates their complex polygenic architecture.

Graphical Abstract



Keywords

Urinary tract stone; genomics; personalized medicine; polygenic risk score; nephrolithiasis; polygenic risk score; genomics

Introduction

Urinary tract stones are highly prevalent in the American population, with up to 10% of men and 7% of women reporting at least one stone during their lifetime.[1] Their composition is

extremely heterogenous with 80% of stones being made up of a mixture of calcium oxalate and calcium phosphate and the majority of the remainder composed of uric acid, struvite and cystine.[2,3] In addition to causing acute discomfort, stones are also associated with increased risk for infection, hydronephrosis, and chronic kidney disease.[3–5] Patients are often diagnosed after presenting with symptoms or incidentally on imaging studies. Current kidney risk prediction tools are limited to recurrence of urinary tract stones rather than prediction of stone formers in the general population.[6,7]

Much progress has been made in understanding the clinical and environmental risk factors predisposing to stone formation. However, these factors do not completely explain the highly heritable nature of urinary tract stones and individuals with no clinical risk factors can still develop urinary tract stones. Prior research has identified several monogenic causes of stones. These monogenic causes are rare, although they are likely underdiagnosed.[8] Although future work may uncover highly pathogenic variants associated with stone formation, it is likely that many urinary stones are of a polygenic nature, due to the combined effect of many genetic factors, each with a small individual effect size.

Genome-wide association studies (GWAS) have generated summary statistics data for millions of single nucleotide variations (SNVs) from large studies and can be used to derive polygenic scores. In other heritable diseases such as schizophrenia and coronary artery disease, these risk score distributions have been applied to identify subgroups (i.e. those with a high polygenic burden) that are at similar risk for developing disease as patients with monogenic mutations, however are much more common.[9–11]. A patient's EHR (electronic health record) is the set of all clinical data available including laboratory measurements, imaging, clinical notes, vitals, diagnostic and procedure codes. Thus the availability of EHR linked biobanks allow for rapid assessment of PRS with many comorbid conditions.

We sought to derive a polygenic score for kidney stones from the UK Biobank and validate it in a large multiethnic biobanked cohort (Bio*Me* Biobank cohort). We tested polygenic scores with increasing number of SNVs. Finally, we assessed whether the best performing polygenic score was associated with urinary tract stones in a subgroup of individuals without clinical risk factors. Finally, we validated this polygenic score in an independent cohort of individuals.

Results

The overall schema of the study is shown in Figure 1.

Study Population

In this study, we included participants with genotyping data (n=28,877) in the Mount Sinai Bio*Me* cohort. We identified 1,071 (4%) participants with a diagnosis of urinary tract stones (cases) and 27,806 participants with no previous diagnosis of urinary tract stones (controls) for all case-control analyses. Participants with urinary tract stones were significantly older (60 vs. 57 years) and had higher BMI (29 vs. 28 kg/m^2) as compared to controls ($p < 0.001$; Table 1). In terms of self-reported race, cases had a significantly greater proportion of Hispanic individuals (44% vs. 32%) and fewer African American (17% vs. 24%) and

European individuals (30% vs. 33%) than controls (p <0.001; Table 1). Cases also had higher proportion of all comorbidities compared to controls (p<0.001; Table 1).

Optimization of PRS

We used summary statistics from an online repository of GWAS results performed on a subset of the UK Biobank cohort of European ancestry (Supplementary Table 1). Using these summary statistics, we varied the rho parameter in LDPred[12] and computed polygenic risk scores for each rho (assumed proportion of casual SNPs)using an additive model. We then selected the PRS corresponding to rho which maximized the association with urinary tract stone diagnosis in Bio*Me*. We observed the greatest association with stone diagnosis for rho =0.001 (Supplementary Figure 1). This score included 7,670,833 SNVs and was used for all downstream analyses and validation.

Association of PRS with Urinary Tract Stone Diagnosis in Bio*Me*

We then computed the the association of the PRS derived from UKBB with urinary tract stone status in Bio*Me*. In order to account for population stratification, we performed our analysis stratified by self-reported race (Figure 2A) and then combined the results with inverse variance weighted meta-analysis. We observed a significant association of PRS with urinary tract stone case/control status in a logistic regression model adjusted for ten genetic principal components (PCs), sex, age, BMI, and traditional risk factors including a history of gout, hypertension, and type 2 diabetes. For every standard deviation (SD) increase in PRS, we observed an increased odds for stone diagnosis of a factor of 1.2 (95% confidence interval 1.1–1.3; p<0.001). (Table 2). The odds of urinary tract stone prevalence increased monotonically as the PRS increased (Figure 2B). The distribution of PRS in Bio*Me* is provided in Figure 2C.

Comparison of genome wide polygenic score to score comprised of only GWAS significant SNPs

In order to assess whether a genome wide polygenic score performs better than one containing GWAS significant SNPs, we then compared the association of the optimized PRS with that of a PRS (PRS_{GWAS}) constructed using the 48 SNPs that met genome wide significance ($P < 5 \times 10^{-8}$) for association with urinary tract stones in UK Biobank. PRS_{GWAS} was not significantly associated with stone diagnosis in Bio*Me* (OR = 0.98; 95% CI: 0.92 – 1.04) adjusted for age, sex, BMI, 10 genetic PCs, and history of gout, hypertension, and Type 2 diabetes.

PRS Associates with Urinary Tract Stones in Absence of Traditional Risk Factors

We then stratified our cohort into a low risk group without any clinical risk factors during their EHR history. Thus, the low risk group did not have any diagnoses codes for hypertension, type 2 diabetes or gout and had an average BMI≤ 25 kg/m^2. 6136 BioMe participants were in the low-risk group (136 stone cases and 6,000 controls) . We then defined high-risk group which had one or more risk factors mentioned above. 22,741 BioMe participants (935 stone cases and 21,806 controls) were in the high-risk group.

In the low risk group, the adjusted odds of stone diagnosis increased by a factor of 1.3 (95% confidence interval 1.1 – 1.6; p = 0.001) per standard deviation of PRS. In the high-risk group, we similarly observed an increased adjusted odds of stone diagnosis of 1.2 (95% confidence interval 1.1 – 1.3; p <0.001) per standard deviation of PRS (Table 3). We observed a monotonic increase in kidney stone diagnosis with PRS in both the low-risk and high-risk groups (Figure 3).

Validation of PRS in External Cohort

We then further validated the PRS in a cohort from the Penn Medicine Biobank (PMBB). The PMBB cohort included 9,973 individuals of European genetic ancestry and and 5,423 individuals of African genetic ancestry (Supplementary Table 2). Since the validation cohort had only a small number of individuals with low clinical risk, we were unable to stratify by clinical risk due to low statistical power. In a logistic regression model adjusted for age, sex, BMI, 10 genetic PCs, gout, hypertension, and type 2 diabetes, we found a significant association between PRS and urinary tract stone diagnosis in European Americans (aOR = 1.16 per SD; 95% CI: 1.06 – 1.27; p <0.001) but not in African-Americans (aOR = 1.02 per SD; 95% CI: 0.89 – 1.16; p=0.07) (Figure 4A, Table 4). However, in a trans-ethnic meta-analysis, we found a significant association with an adjusted odds ratio of 1.12 (95% CI: 1.05 – 1.2; p<0.001) for every SD increase in PRS (Table 4). Prevalence of stone diagnosis increased monotonically as a function of PRS in European Americans (Figure 4B) but not in African Americans (Figure 4C).

Improvement in Risk Discrimination with addition of PRS

We then assessed the improvement in discriminative ability of the PRS in Bio*Me* (since PMBB had only high-risk individuals). The bootstrapped AUC (1000 iterations) of a logistic regression model incorporating only age, sex, and clinical covariates was 0.63 (95% CI: 0.62–0.65). After adding PRS, AUC improved to 0.68 (95% CI: 0.66–0.70). This represents a modest, but significant improvement in prediction.

Discussion

Utilizing the combined availability of linked genotypic and clinical diagnosis data, we developed and validated a genome-wide polygenic risk score for urinary tract stones in two independent cohorts.

The advent of GWAS and large-scale population-based genetic studies such as the UK Biobank[13,14] have led to the identification of the polygenic nature of complex disorders such as hypertension[15], coronary artery disease[11], stroke[16], and schizophrenia[17]. Since the genome does not change over one's lifetime, polygenicity can be leveraged for better understanding the genetic architecture of complex disease. In this study, we demonstrate that using genome-wide information summarized as a polygenic score significantly correlates with a stone diagnosis. We also show that individuals with a PRS in the top 10th are at almost three-fold odds of a urinary tract stone diagnosis as compared to individuals in the bottom 10th percentile of PRS, indicating these individuals are at higher baseline genetic risk.

Urinary tract stone formation is a complex process with known risk factors, including family history, obesity, gout, type 2 diabetes as well as dietary factors such as excessive animal protein and salt intake and insufficient water intake[6]. In comparing cases with controls, we recapitulated known epidemiologic associations, with greater incidence of obesity, hypertension, type 2 diabetes, gout and proportion of males in cases[18–21]However, we also demonstrate that our genome-wide polygenic risk score is significantly associated with urinary stone diagnosis in low risk individuals with no obesity, hypertension, type 2 diabetes, or gout. While this is not a complete set of clinical risk factors and does not consider family history, it represents the first attempt at applying polygenic scores in individuals with no evident clinical risk factors. This indicates that urinary tract stones may have a genetic component that is additive to known clinical risk.

We also validated the PRS in an external cohort and show that the PRS is associated with stone diagnosis in European Americans but not African Americans. This failure to replicate in African Americans is likely because the GWAS summary statistics used to generate the PRS were derived from a European cohort. This highlights a crucial equity issue. There is underrepresentation of ethnic minorities in genomic research and thus a need to conduct genomic analyses in multiethnic datasets[22]. Although we cannot make claims about clinical utility, with the increasing availability of genetic information linked to clinical phenotype information, polygenic scores could be utilized to identify individuals, especially those without traditional risk factors with a high risk of developing urinary tract stones.[23] However, a polygenic score should be considered as a biomarker, albeit constant and further studies using clinical utility of these scores need to be explored further to determine the feasibility of incorporating polygenic scores in clinical practice. “When the PRS was added to a logistic regression model with only clinical variables, the AUC for urinary tract stone prediction increased by 5%. This represents a moderate increase in discriminative ability. However, since family history and diet were not available in this study, these variables must be considered as covariates when evaluating the added predictive ability of the PRS.”

This study should be interpreted in the light of some limitations. First, we defined stone diagnosis utilizing ICD codes. Although ICD codes have been previously used in epidemiological studies[24,25] and Semins et al. previously reported a high positive predictive value of 96% for using ICD codes to identify urinary tract stones in EHR systems[26], the possibility of misclassification exists. The overall prevalence of urinary tract stone diagnosis was 3.7% in Bio*Me* and 5.6% in PMBB, which is significantly lower than the population prevalence estimates of approximately 10%. Second, our study is limited by the scope of EHR data. Since EHR data is intended to serve a billing tool rather than a complete patient phenotype, we did not have access to records of dietary habits, environmental influences, or family history of urinary tract stones. Thus, we are unable to determine the relative predictive value of these non-genetic factors. However, even if a subset of the low risk patients has a family history of kidney stones, our polygenic risk score can still be robustly applied in this setting since many patients may be unaware of their family history. In this case, applying the polygenic risk score would provide an unbiased estimate of kidney stone risk. While lack of complete phenotypic information is a restriction to using EHR data, by using this cohort, our approach shows utility of this approach in large multiethnic populations. Third, our validation cohort (PMBB) included only a small number of

individuals with low clinical risk due to differences in recruitment strategies. Thus, due to low statistical power, we could not conduct stratification based on clinical risk like in the Bio*Me* Biobank. Fourth, since our dataset is retrospective, we are unable to determine the clinical utility of a PRS as a screening tool to guide clinical management. Fifth, since we use millions of SNPs with small but non-zero effect sizes and analysis uses linear regression on a binary outcome, it is difficult to interpret the contribution of individual variants in isolation. However this emphasizes the polygenic architecture of many complex diseases. Finally, we could not assess whether PRS associates with number, size, type, recurrence and composition of kidney stones, since this information is not easily accessible in routinely collected EHR data. However, these scores can be validated in other existing cohorts of kidney stone patients with granular information about kidney stones.

In summary, we show that a genome-wide polygenic score derived from publicly available GWAS data and validated within two large, multiethnic population biobanks is significantly associated with urinary stone diagnosis. We show that in the absence of clinical risk factors, the polygenic risk score was significantly associated with urinary tract stones, Although further validation and implementation needs to be conducted in prospective longitudinal studies, this work highlights the highly polygenic nature of this complex disease.

Conclusions

A genome-wide polygenic risk score is significantly associated with urinary tract stone diagnosis independent of clinical risk factors.

Methods

Derivation of Polygenic Scores

Polygenic scores measure the cumulative impact of common variants on the risk of certain disorders with the impact of individual variants assumed to be additive. Thus, for each individual, scores are computed by taking the sum of the dosage of risk variants weighted by their effect on the disease under consideration.

We derived all polygenic scores using data from the UK Biobank. Briefly, the UK Biobank consists of genotype, phenotype, and demographic data of more than 500,000 individuals recruited across the United Kingdom enrolled between ages 40 and 69. Each individual completed a questionnaire with medical history and demographics. Additionally, blood biochemistry assays were performed at time of enrollment. Individual genotypes were generated from either the Affymetrix Axiom UK Biobank array (~450,000 individuals) or the UK BiLEVE array (~50,000 individuals), each containing ~0.8 million markers. More variants were then imputed using the Haplotype Reference Consortium (HRC) combined with UK10K haplotype resource, leading to ~96 million variants available.

We obtained genome-wide association study (GWAS) summary statistics from an online GWAS UK Biobank database[27] for the kidney stone/ureter stone/bladder stone phenotype. This phenotype was identified using diagnostic codes. The summary statistics were generated by first restricting SNPs to those with minor allele frequency >0.1% and Hardy-

Weinberg equilibrium p-value $>1\times10^{-10}$. Following quality control, 10.8 million SNPs were included in downstream association analysis. Association analysis was conducted by fitting a logistic regression model with additive genotype coding (0,1, or 2 copies of the minor allele) adjusted for sex and the first ten genetic principal components.

To compute polygenic scores, we first adjusted summary statistics effect sizes using LDpred[12]. This tool adjusts effects sizes for linkage disequilibrium. We used the 1000 Genomes European population LD reference panel. The key tuning parameter in LDpred is rho, the assumed proportion of causal SNPs. We used effect sizes adjusted for several values of rho between 1 (assumes all SNPs are causal) and 0.001. We then computed a polygenic score for each using an additive function such that $PRS=\sum iSi\times Gi$, where Si = adjusted beta statistics for minor allele and Gi = genotype (0,1, or 2). We chose an optimal rho by selecting the rho and the PRS that had a maximal association with urinary tract stone diagnosis in Bio*Me*. We scaled the polygenic scores to have mean zero and standard deviation of one.

A score was computed with only GWAS significant SNPs that met the P value threshold ($P <5 \times 10^{-8}$) for association with urinary tract stones in UK Biobank. Weights from the GWAS summary statistics were directly used without LDpred adjustment.

Bio*Me* Discovery Cohort

We utilized the Bio*Me* Biobank at Mount Sinai. Briefly, the Bio*Me* Biobank is an electronic health records (EHR)-linked clinical care cohort comprised of over 45,000 participants from diverse ancestries (African, Hispanic/Latino, European and Other ancestries), with accompanying genome-wide genotyping data for 31,441 participants. Along with the genetic information, Bio*Me* is linked to a wide array of biomedical traits, originating from Mount Sinai's system wide Epic EHR. Enrollment of participants is predominantly through ambulatory care practices and is representative of Mount Sinai's larger patient population. BioMe participants (N=31,441) were genotyped on the Illumina Global Screening Array (GSA) platform. Quality control and imputation of the GSA data is detailed in Supplementary methods.

Penn Medicine BioBank (PMBB) Validation Cohort

The PMBB consists of 60,000 patients recruited from clinical sites across the University of Pennsylvania Health System who have provided consent for access to all electronic health records. This study included a subset of 9,973 European and 5,423 African ancestry patients that had undergone genotyping. Samples within PMBB were genotyped on the Illumina Quad Omni SNP Chip by Regeneron. Following sequencing, standard QC procedures were followed to remove rare and missing variants as well as variants in linkage disequilibrium. Variants with a minor allele frequency of < 0.05, a missing rate of > 0.05 and a Hardy-Weinberg P of $> 10^6$ were removed. Furthermore, samples with a genotype missingness of > 0.02 were also removed.

Identification of Urinary Tract Stone Patients

We defined urinary tract stone cases in Bio*Me* and PMBB as any patient with one or more encounters with an international classification of diseases-Clinical Modification 10(ICD-10) diagnosis code of N20.0, N20.1, N20.2, or N20.9 or ICD-9 diagnosis code of 592.0, 592.1, or 592.9.

Statistical Analysis

We compared categorical and continuous variables between cases and controls using the Chi-squared test and t-test, respectively. To test for association between scaled PRS and urinary tract stone diagnosis, we utilized a logistic regression model adjusting for age, sex, ten principal components (PCs), and history of clinical comorbidities, gout, hypertension, and Type 2 diabetes determined using previously validated phenotyping algorithms[28] and diagnostic codes. For all association analyses, we stratified participants by self-reported race/ethnicity (European American, African American, Hispanic American, and Other). Using race-specific results, we performed a meta-analysis using the inverse variance method as implemented in the *meta* R package[29].

We then sought to assess the utility of the best performing polygenic score in improving discrimination of stone diagnosis in the absence and presence of clinical risk factors, including gout, hypertension, type 2 diabetes and body mass index (BMI). A low risk group was defined as individuals with BMI <25 and no history of hypertension, type 2 diabetes, or gout. The high-risk group included all other individuals. We then fit a logistic regression model adjusted for age, sex, and 10 genetic PCs to determine the association between polygenic scores and kidney stone diagnosis in each subgroup separately stratified by race. We also used a bootstrap method with 1000 iterations to determine the area under the receiver operating curve (AUC) of logistic regression models with 1) clinical variables and 2) clinical variables + PRS. 95% confidence intervals of AUC were computed from bootstrap results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The BioMe healthcare delivery cohort at Mount Sinai was established and maintained with a generous gift from the Andrea and Charles Bronfman Philanthropies. This work was supported in part through the computational resources and staff expertise provided by the Department of Scientific Computing at the Icahn School of Medicine at Mount Sinai.

Funding

GNN is supported by a career development award from the National Institutes of Health (NIH) (K23DK107908) and is also supported by R01DK108803, U01HG007278, U01HG009610, and 1U01DK116100. SGC is also supported by the following grants: R01DK106085, R01HL85757, R01DK112258, and U01OH011326. RD is supported by R35GM124836 from the National Institute Of General Medical Sciences of the National Institutes of Health and R01HL139865 from the National Heart, Lung, Blood Institute of the National Institutes of Health and previously an American Heart Association Cardiovascular Genome-Phenome Discovery grant (15CVGPSD27130014). PQD was supported by NIH Medical Scientist Training Program Training Grant (T32GM007205). SMD is supported by the US Department of Veteran Affairs (IK2-CX001780).

References

1. Scales CD, Smith AC, Hanley JM & Saigal CS Prevalence of kidney stones in the United States. Eur. Urol (2012). doi:10.1016/j.eururo.2012.03.052
2. Evan AP Physiopathology and etiology of stone formation in the kidney and the urinary tract. Pediatric Nephrology (2010). doi:10.1007/s00467-009-1116-y
3. Khan SR et al. Kidney stones. Nat. Rev. Dis. Prim 2, 16008 (2016). [PubMed: 27188687]
4. Rule AD et al. Kidney stones and the risk for chronic kidney disease. Clin. J. Am. Soc. Nephrol (2009). doi:10.2215/CJN.05811108
5. Monico CG & Milliner DS Genetic determinants of urolithiasis. Nature Reviews Nephrology (2012). doi:10.1038/nrneph.2011.211
6. Alelign T & Petros B Kidney Stone Disease: An Update on Current Concepts. Adv. Urol (2018). doi:10.1155/2018/3068365
7. Punnoose AR, Golub RM & Lynm C Kidney Stones. JAMA 307, 2557–2557 (2012). [PubMed: 22797461]
8. Goldfarb DS The Search for Monogenic Causes of Kidney Stones. J. Am. Soc. Nephrol (2014). doi:10.1681/asn.2014090847
9. Khera AV et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nature Genetics (2018). doi:10.1038/s41588-018-0183-z
10. Fanous AH et al. Genome-wide association study of clinical dimensions of schizophrenia: Polygenic effect on disorganized symptoms. Am. J. Psychiatry (2012). doi:10.1176/appi.ajp.2012.12020218
11. Inouye M et al. Genomic Risk Prediction of Coronary Artery Disease in 480,000 Adults: Implications for Primary Prevention. J. Am. Coll. Cardiol (2018). doi:10.1016/j.jacc.2018.07.079
12. Vilhjálmsson BJ et al. Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. Am. J. Hum. Genet (2015). doi:10.1016/j.ajhg.2015.09.001
13. Bycroft C et al. Genome-wide genetic data on ~500,000 Biobank participants. bioRxiv (2017). doi:10.1101/166298
14. Thompson SG & Willeit P UK Biobank comes of age. Lancet (2015). doi:10.1016/s0140-6736(15)60578-5
15. Kraja AT et al. New Blood Pressure-Associated Loci Identified in Meta-Analyses of 475 000 Individuals. Circ. Cardiovasc. Genet (2017). doi:10.1161/CIRCGENETICS.117.001778
16. Pulit SL et al. Atrial fibrillation genetic risk differentiates cardioembolic stroke from other stroke subtypes. Neurol. Genet (2018). doi:10.1212/NXG.0000000000000293
17. Ward J et al. Genome-wide analysis in UK Biobank identifies four loci associated with mood instability and genetic correlation with major depressive disorder, anxiety disorder and schizophrenia. Transl. Psychiatry (2017). doi:10.1038/s41398-017-0012-7
18. Taylor EN, Stampfer MJ & Curhan GC Diabetes mellitus and the risk of nephrolithiasis. Kidney Int. (2005). doi:10.1111/j.1523-1755.2005.00516.x
19. Cappuccio FP, Strazzullo P & Mancini M Kidney stones and hypertension: population based study of an independent clinical association. BMJ (1990).
20. Clayman RV Obesity, Weight Gain, and the Risk of Kidney Stones. J. Urol (2005). doi:10.1016/s0022-5347(01)68973-0
21. Yu TF Urolithiasis in hyperuricemia and gout. Journal of Urology (1981). doi:10.1016/S0022-5347(17)54561-9
22. Popejoy AB & Fullerton SM Genomics is failing on diversity. Nature (2016). doi:10.1038/538161a
23. Frassetto L & Kohlstadt I Treatment and prevention of kidney stones: An Update. Am. Fam. Physician (2011).
24. Tasian GE et al. Oral Antibiotic Exposure and Kidney Stone Disease. J. Am. Soc. Nephrol (2018). doi:10.1681/asn.2017111213
25. Assimos D Kidney stones associate with increased risk for myocardial infarction. Journal of Urology (2011). doi:10.1016/j.juro.2011.01.049

26. Semins MJ, Trock BJ & Matlaga BR Validity of Administrative Coding in Identifying Patients With Upper Urinary Tract Calculi. J. Urol (2010). doi:10.1016/j.juro.2010.03.011
27. Neale B Rapid GWAS of Thousands of Phenotypes for 337,000 Samples in the UK Biobank.
28. Nadkarni GN et al. Development and validation of an electronic phenotyping algorithm for chronic kidney disease. AMIA … Annu. Symp. proceedings. AMIA Symp (2014).
29. Schwarzer G, Carpenter JR & Rücker G An Introduction to Meta-Analysis in R. in (2015). doi:10.1007/978-3-319-21416-0_1

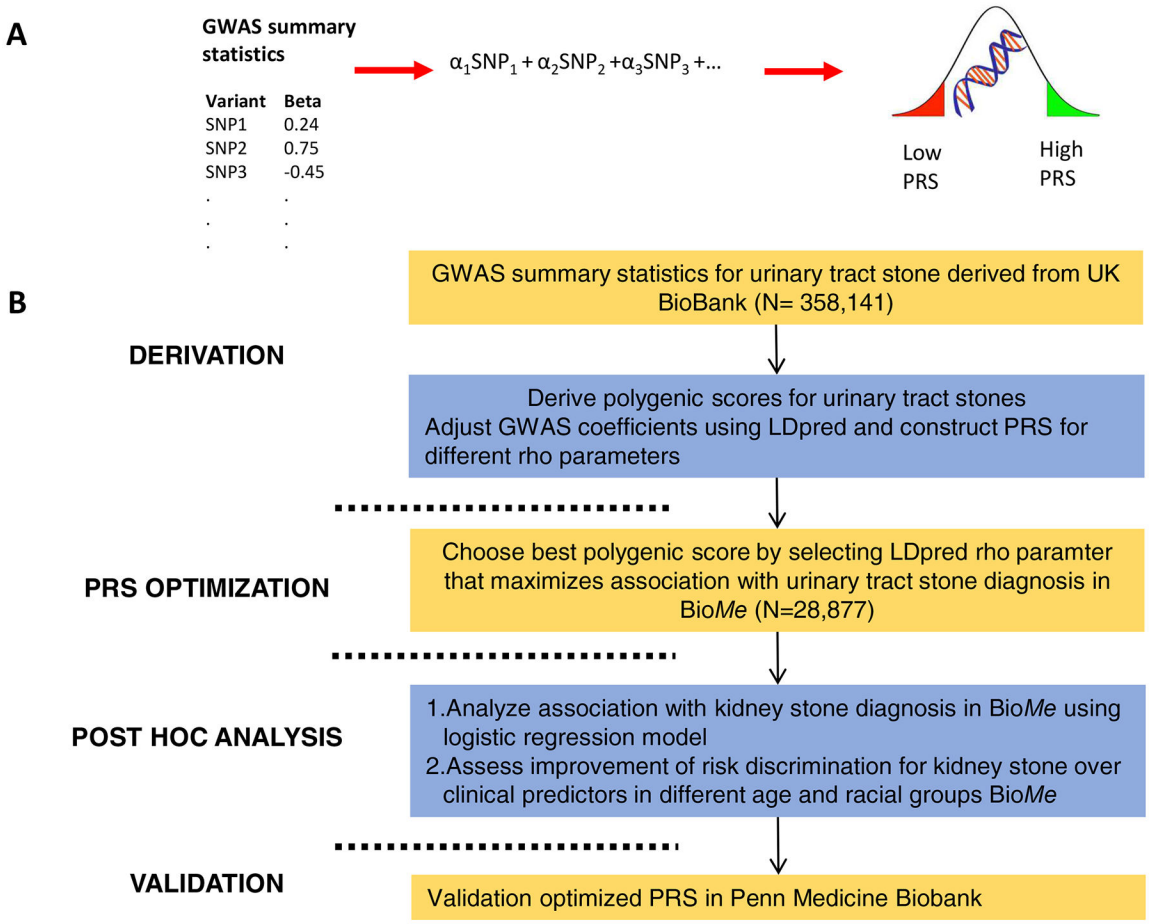


Figure 1.

A. Generation of polygenic risk score from GWAS summary statistics**. B.** Overall scheme of polygenic risk score development and application in BioMe and PMBB cohorts

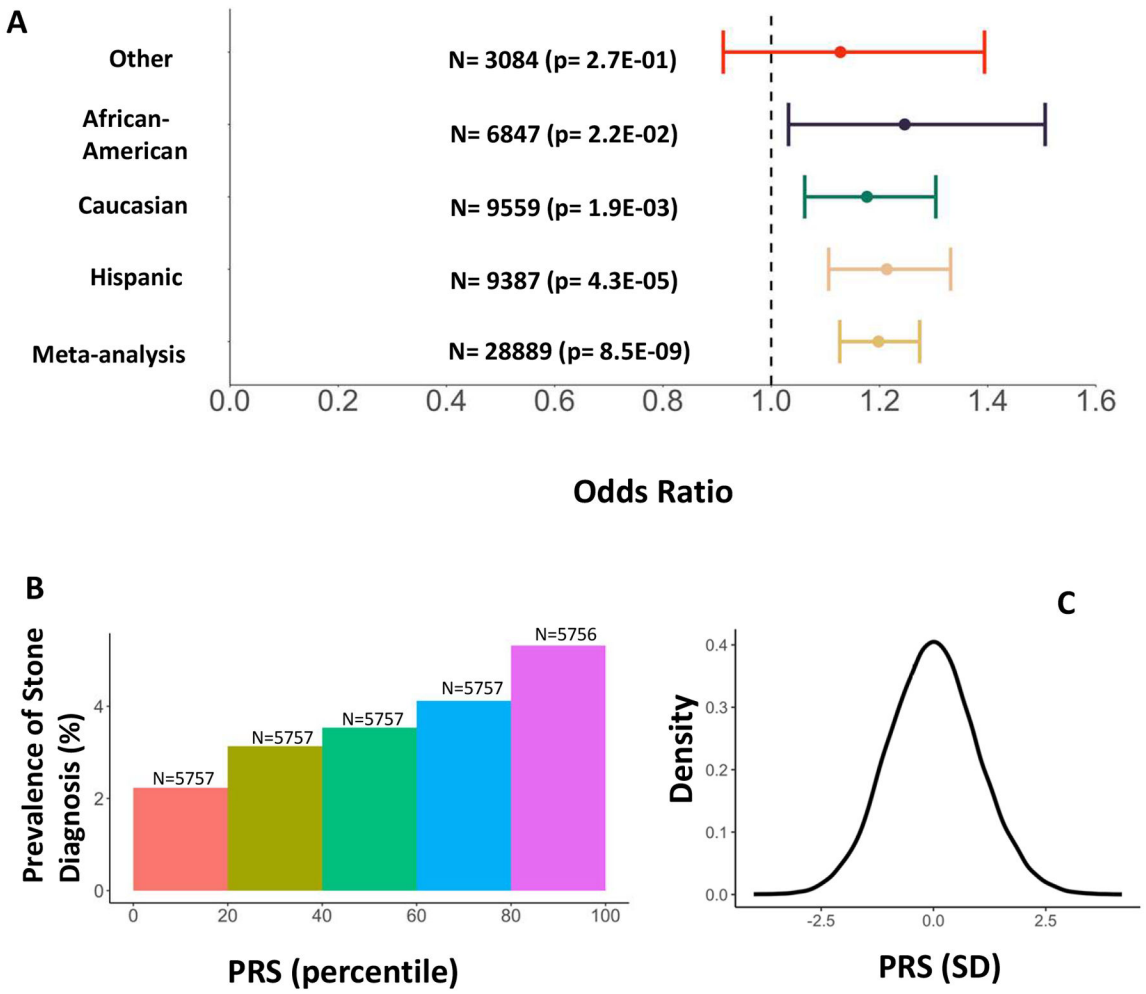


Figure 2.

A. Association of PRS with urinary tract stone diagnosis stratified by race in Bio*Me*. Odds ratio was computed using a logistic regression adjusted for age, sex, 10 genetic PCs, history of hypertension, gout, and type 2 diabetes. **B.** Prevalence of urinary tract stone diagnosis for different levels of PRS **C.** Distribution of normalized PRS in Bio*Me* .

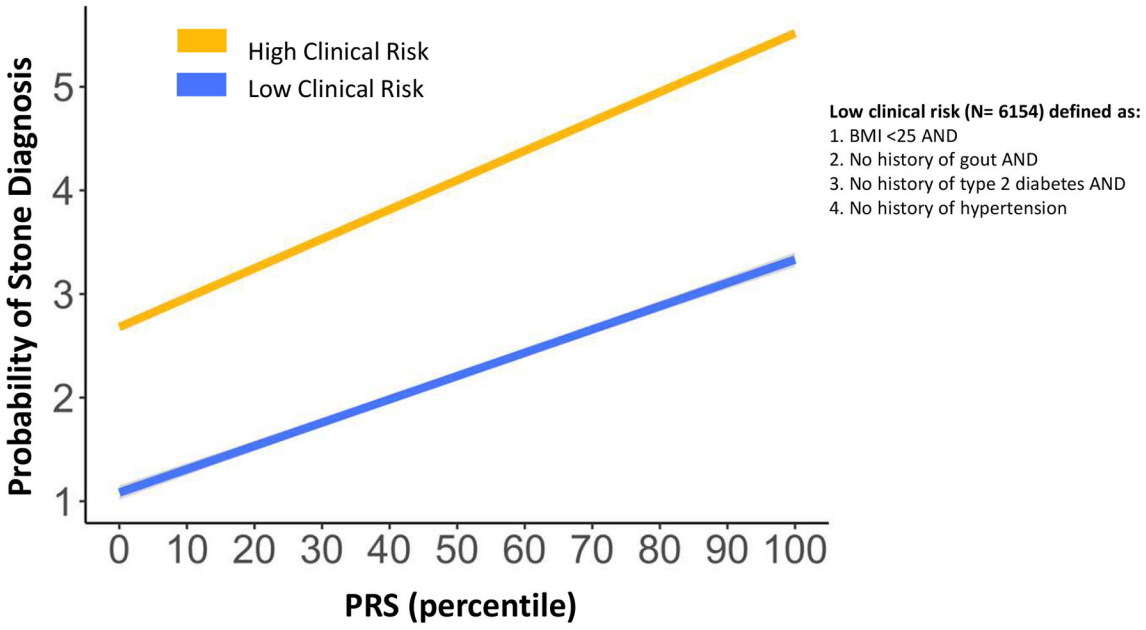


Figure 3.

Association of PRS with urinary tract stone diagnosis in Bio*Me* stratified by clinical risk adjusted for age, sex, BMI, and 10 genetic principal components. Low clinical risk was defined as BMI <25 and no history of gout, type 2 diabetes and hypertension. High clinical risk included all other patients.

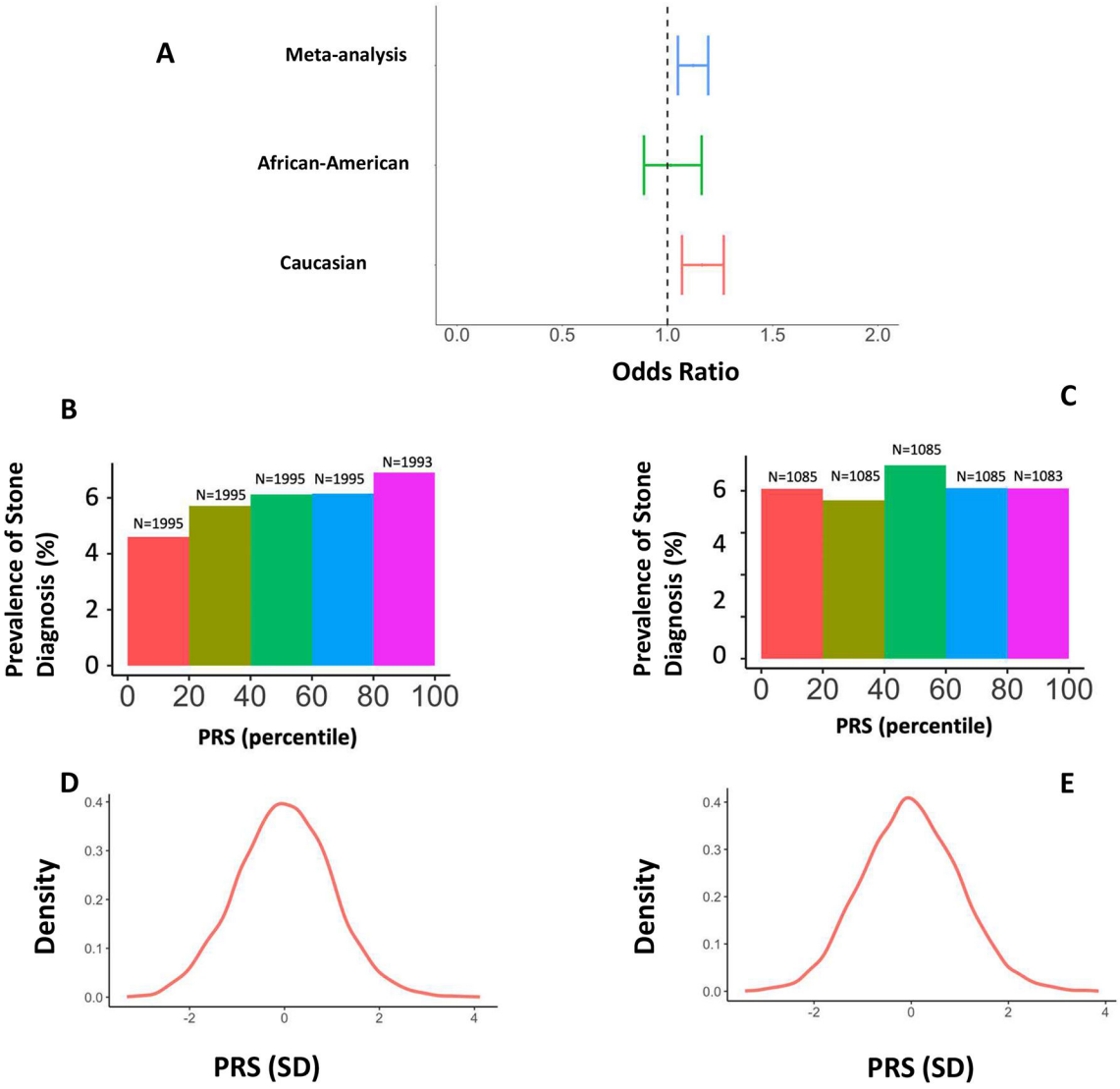


Figure 4.

A. Association of PRS with urinary tract stone diagnosis in Penn Medicine Bio Bank validation cohort stratified by self-reported race. Prevalence of urinary tract stone diagnosis for different levels of PRS in **B.** European Americans and **C.** African Americans. Distribution of PRS in **D.** European Americans and **E.** African Americans

Table 1.

Baseline Demographic and Clinical Characteristics of Urinary Tract Stone Cases and Controls

	Case (N=1071)	**Control (N=27806)**	**P**
Male, n (%)	526 (49%)	11825 (43%)	<0.001
Age, Mean (SD)	60 (14)	57 (18)	<0.001
Race, n (%)			
African American	180 (17%)	6670 (24%)	<0.001
European American	323 (30%)	9248 (33%)	
Hispanic American	473 (44%)	8918 (32%)	
Other	95 (9%)	2992 (10.8%)	
Body Mass Index in kg/m^2, Mean (SD)	29 (7)	28 (7)	<0.001
Clinical Comorbidities			
Hypertension, n(%)	723 (68%)	13147 (47%)	<0.001
Coronary Artery Disease, n(%)	345 (31%)	5980 (22%)	<0.001
Type 2 Diabetes, n(%)	341 (34%)	5714 (21%)	<0.001
Obesity, n (%)	415 (39%)	8314 (30%)	<0.001
Gout, n (%)	64 (6%)	744 (3%)	<0.001

Table 2

Association of PRS with urinary tract stone diagnosis adjusted for age, sex, BMI, 10 genetic PCs, history of hypertension, gout and type 2 diabetes stratified by race.

Group	Number of Cases	Number of Controls	Adjusted Odds Ratio (95% Confidence Interval)	P
European American	323	9248	1.18 (1.06–1.3)	0.002
Hispanic American	473	8918	1.21 (1.11–1.33)	<0.001
African American	180	6670	1.25 (1.03–1.51)	0.02
Other	95	2992	1.13 (0.91–1.39)	0.2
Meta-analysis	**1071**	**27828**	**1.20 (1.13–1.27)**	<0.001

Table 3

Association of PRS with urinary tract stone diagnosis in clinically high risk and low risk subgroups in Bio*Me* adjusted for age, sex, BMI, and 10 genetic PCs.

High Clinical Risk					
Group	**P value**	**Odds Ratio**	**95 % Confidence Interval**	**Number of Cases**	**Number of Controls**
Hispanic American	0.002	1.17	(1.06, 1.29)	428	7753
African American	0.07	1.20	(0.98, 1.45)	166	5861
European American	0.003	1.19	(1.06, 1.33)	265	6151
Other	0.3	1.12	(0.89, 1.4)	76	2041
Meta-analysis	**<0.001**	**1.18**	**(1.10, 1.26)**	**935**	**21806**
Low Clinical Risk					
Hispanic American	0.0004	1.72	(1.27, 2.32)	45	1161
African American	0.096	1.84	(0.91, 3.87)	14	806
European American	0.3	1.13	(0.88, 1.44)	58	3085
Other	0.4	1.24	(0.72, 2.18)	19	948
Meta-analysis	0.001	1.33	(1.12, 1.58)	136	6000

Table 4

Association of PRS with urinary tract stone diagnosis in external validation cohort (Penn Medicine Biobank). A logistic regression model adjusted for age, sex, BMI, and 10 genetic principal components was fit in racial groups. A meta-analysis was performed using the inverse variance method

Group	Number of Cases	Number of Controls	Odds Ratio (95% Confidence Interval)	P value
African American	223	5200	1.02 (0.89–1.16)	0.07
European American	588	9385	1.16 (1.07–1.27)	<0.001
Meta-analysis	**811**	**14585**	**1.12 (1.05–1.19)**	**<0.001**