

Original

Polygenic risk of bone fractures in Spanish women with osteoporosis

Álvaro del Real¹, Raquel Cruz², José M. Olmos^{1,3}, José L. Hernández^{1,3}, Carmen Valero^{1,3}, José A. Riancho^{1,3}

¹Department of Medicine and Psychiatry. Universidad de Cantabria. IDIVAL. Santander, Spain. ²Grupo de Medicina Xenómica. Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER). Universidad de Santiago de Compostela. Santiago de Compostela, A Coruña. Spain. ³Department of Internal Medicine. Hospital Universitario Marqués de Valdecilla. IDIVAL. Santander, Spain

Abstract

Background: osteoporosis is a highly polygenic trait characterized by low bone mineral density (BMD) and/or fragility fractures. Over the past decade, polygenic risk scores (PRS) are an emerging tool to try to predict the risk of complex disorders with a genetic component.

Objective: to analyze the capacity of different PRSs to predict osteoporosis in the Spanish population.

Material and methods: our dataset consisted of two differentiated groups. The first group included osteoporosis cases diagnosed and treated at the Marques de Valdecilla University Hospital ($n = 304$; 293 women) while the second group consisted of people from the overall Spanish population ($n = 3199$; 1458 women). Four previously generated PRSs were compared with generalized linear models.

Results: the osteoporosis group showed a significantly higher genetic risk compared to the control group in 3 PRSs (PRS-1 $p = 1e-7$; PRS-2 $p = 1.87e-15$; PRS-3 $p = 0.1477$; PRS-4 $p = 8.98e-9$). In addition, in these PRSs, the individuals in the upper quartile of risk had a significantly higher risk of osteoporosis, compared to those individuals in the other quartiles (PRS-1 OR, 1.83; PRS-2 OR, 2.11; PRS-3 OR, 0.96; PRS-4 OR, 1.72).

Conclusions: in summary, the application of PRSs shows significant differences between the overall Spanish population and patients with osteoporosis, which is suggestive of its utility within strategies for the identification of subjects at risk based on clinical-genetic criteria.

Keywords:
Osteoporosis.
Polygenic risk
scores. Genomics.
Bone mineral
density.

Received: 19/10/2022 • Accepted: 19/04/2023

Acknowledgements: partially supported by a grant from Instituto de Salud Carlos III (PI21/00532). Alvaro del Real received support by the postdoctoral grant Margarita Salas in the University of Cantabria, Spain.

Conflicts of Interest: the authors declare no conflict of interest.

Del Real Á, Cruz R, Olmos JM, Hernández JL, Valero C, Riancho JA. Polygenic risk of bone fractures in Spanish women with osteoporosis. *Rev Osteoporos Metab Miner* 2023;15(2):66-71

DOI: 10.20960/RevOsteoporosMetabMiner.00014

Correspondence:

Álvaro del Real. Department of Medicine and Psychiatry. Universidad de Cantabria. IDIVAL. Avda. Cardenal Herrera Oria, s/n. 39011 Santander, Spain
e-mail: delreal@unican.es

INTRODUCTION

Osteoporosis is the most prevalent bone disease characterized by low bone mineral density (BMD) leading to an increased risk of fracture. Osteoporosis-related fractures represent an immense economic burden on the healthcare systems. Common diseases such as osteoporosis are usually polygenic, involving many genetic variants rather than rare monogenic mutations (1). Several studies have shown that BMD is a highly polygenic trait, which is directly associated with bone fracture (2). Over the past 15 years, genome-wide association studies (GWAS) have identified many genomic loci as related to the risk of various complex diseases (3). The knowledge of the genetic variants involved in a specific trait allows the early identification of subjects at risk and the initiation of preventive measures. Thus, in the GWAS era, several genetic variants have been related to BMD and fracture risk. The seminal study of Estrada et al. identified 56 loci associated with BMD and 13 SNPs with bone fractures (4). Another more recent study from the UK Biobank database analyzed the association of genetic variants and heel quantitative ultrasound (eBMD) in a total of 426 824 individuals to finally identify 518 significant loci (5).

The results of those studies are being used to develop risk scores based on the analysis of multiple gene variants (polygenic risk scores, PRSs). Hence, a PRS can be defined as an individual's mark made of the allelic signature at a number of polymorphic loci (often, tens or hundreds, or even thousands) related to the genetic susceptibility to develop a disorder (2). A representative analysis with PRSs involves an association between a PRS and a trait from the main data. This association can be evaluated with standard analytical procedures such as the *p* value to test a null hypothesis; effect size estimate (OR of high vs low risk individuals), and/or with measures of discrimination like the area under the curve (AUC). Several statistical tests can be applied to check the significance of the association including linear or logistic regression with or without adjusting for covariates like sex and age (6).

Several PRSs have been generated in relation to BMD and/or risk of fracture. First, the 56 loci identified (*n* = 63 probes) from the seminal study described above (4) were used as a PRS related to femoral neck BMD. Richards' lab developed a prediction of fracture risk PRS (*n* = 21717 probes) by using ultrasonography data of the calcaneus as an intermediate phenotype (7). Their polygenic risk score was more strongly associated with the risk of fracture than many other clinical risk factors, including age, sex, BMI and FRAX clinical factors (8). Additionally, Tanigawa et al. generated two distinct PRS models, incorporating the results at 316 and 1270 loci, respectively (9).

The objective of this study was to analyze the capacity of the previously mentioned PRSs to discriminate between patients with osteoporosis and controls in the Spanish population.

METHODS

SAMPLE RECRUITMENT

Our dataset included two groups. The group of cases corresponds to patients with osteoporosis recruited at the Marques de Valdecilla University Hospital (*n* = 304; 293 women; mean age, 65 years; range 47 to 87 years). Subjects with secondary osteoporosis were excluded. BMD was measured by dual X-ray densitometry (DXA) at the spine (mean BMD 0.744 [Interquartile range, IR, 0.692 to 0.792]) and the hip (mean BMD 0.737 [IR 0.679 to 0.803]) using a Hologic QDR 4500 densitometer (Waltham, MA, United States).

The control included samples from the overall Spanish population that were provided by the "Banco Nacional de ADN Carlos III (BNADN; www.bancoadn.org) (*n* = 3199; 1458 women; mean age, 48 years; range 18 to 104 years).

These data were not included in former GWAS. The study protocol was approved by the institutional review board (Comité de Ética en Investigación Clínica de Cantabria). All patients gave their informed written consent.

DNA ISOLATION AND GENOTYPING

DNA was isolated from aliquots of peripheral blood using commercially available column-based kits, following the manufacturer's instructions for use. Quantification of DNA was performed using Qubit dsDNA BR Assay Kit (ThermoFisher, Waltham, MA, United States). DNA samples from both groups were genotyped at the Spanish National Genotyping Center ("Centro Nacional de Genotipado-Fundación Pública Galega de Medicina Xenómica"), using the Axiom™ Spain Biobank array following the manufacturer's instructions for use (Axiom™ 2.0 Assay 96-Array Format Manual Workflow; ThermoFisher Scientific). Briefly, total genomic DNA (200 ng) was amplified and randomly fragmented into 25 to 125 base pair fragments, which were then purified and resuspended in a hybridization cocktail. The hybridization-ready targets were then transferred to the GeneTitan Multichannel Instrument for automated, hands-free processing (including hybridization to Axiom array plates, staining, washing and imaging). CEL files were automatically processed for allele calling using the Axiom GT1 algorithm available through the Axiom Analysis Suite v4.0.3.3 and following the Axiom™ Genotyping Solution Data Analysis User Guide (ThermoFisher Scientific, Waltham, MA, United States).

QUALITY CONTROL AND PRS ANALYSIS

The Axiom Analysis Suite was applied to conduct the quality control of genotyped data. Thresholds applied were DQC \geq 0.85, and call rate \geq 97 %. The percent of

passing samples was ≥ 95 , and the average call rate for passing samples was ≥ 98.5 . To assess the existence of stratification and to identify kinship relationships, PCA and IBD analyses were implemented with PLINK software. After that, genotyped data was imputed with TOPMED imputation software. Minor allele frequency (MAF) or Hardy-Weinberg equilibrium (HWE) thresholds were not implemented because of the possibility of losing selected probes from the PRSs databases.

Four different PRSs datasets were computed based upon previous publications. Details for each PRS are shown on table I.

Statistical tests were performed with the R software (version 4.2.1). Moreover, ROC curves were generated with the "pROC" package (10).

RESULTS

Three of the 4 PRSs showed significantly higher scores in the osteoporosis group compared to the control group (PRS-1 $p = 1e-7$; PRS-2 $p = 1.87e-15$; PRS-3 $p = 0.1477$; PRS-4 $p = 8.98e-9$). Moreover, in those three PRSs with significant differences the individuals from the risk quartile (which corresponds to the first quartile for PRS-2 and upper quartile for PRS-1 and PRS-4) had a significantly higher risk of osteoporosis compared to those individuals from the other quartiles (PRS-1 OR, 1.83 [CI, 1.41-2.36]; PRS-2 OR, 2.11 [CI, 1.64-2.71]; PRS-3 OR, 0.96 [CI, 0.72-1.27]; PRS-4 OR, 1.72 [CI, 1.32-2.21]) (Fig. 1).

Furthermore, the frequencies and the probability of disease are shown on figure 2. PRS-1 and PRS-4 have

Table I. Details of the published PRSs used in this article

	Associated phenotype	Probes	Merged probes	Ancestry	Reference
PRS_1	Femur neck BMD	63	63	Mainly European and partly East Asian	Estrada K, et al. Nat Genet 2012;44:491-501
PRS_2	Heel quantitative speed of sound (SOS)	21716	15 721	Predominantly white British for training, testing and validation sets	Forgetta V, et al. PLOS Medicine 2020;17(7):e1003152
PRS_3	Osteoporosis	316	273	White British for training, testing and validation sets.	Tanigawa Y, et al. PLoS Genet 2022;18(3):e1010105
PRS_4	Osteoporosis without fracture	1270	1136		

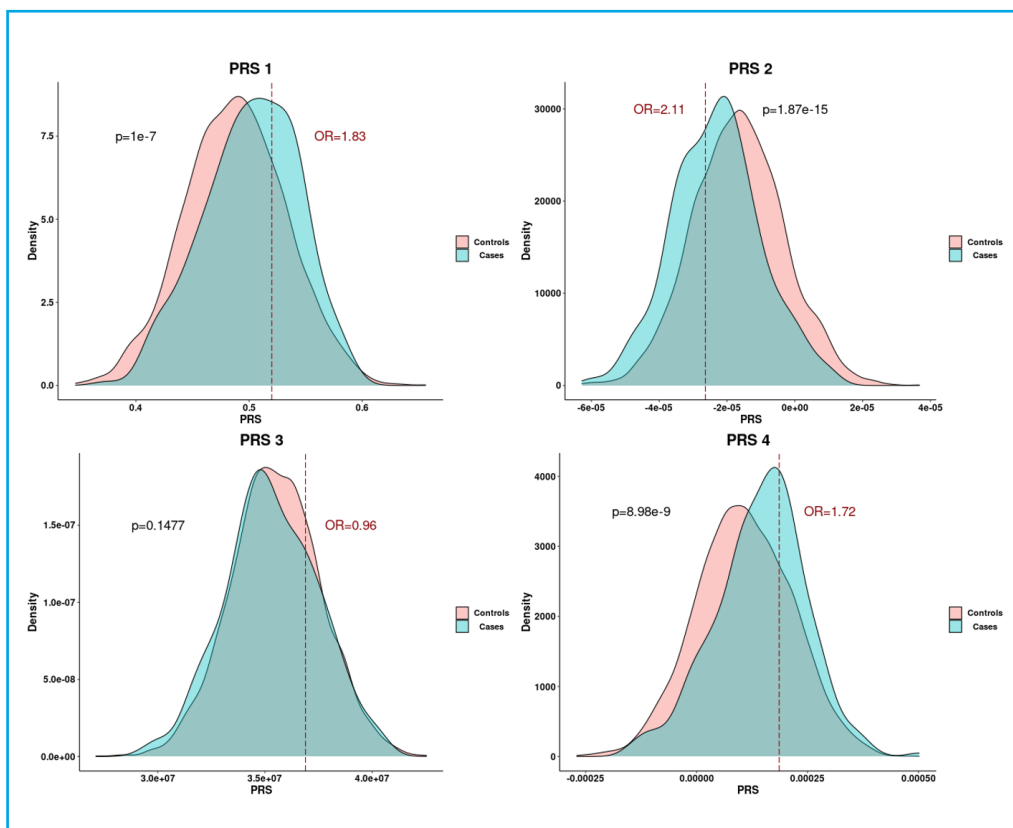


Figure 1. Density plots with osteoporotic cases in blue and controls in orange for each PRS tested in the manuscript. The p value is estimated using Student t tests to look for the differences in the mean of each group. Vertical red dot line limits the upper quartile (first quartile for PRS-2), which should have a higher risk of cases. The quartiles are calculated with all samples together, not by group. Moreover, Odds ratio (OR) are estimated from the quartile selected vs the other quartiles.

an ascending curve because the higher the score the higher the risk of osteoporosis. PRS-2 values are inversely correlated to osteoporosis because it is associated with heel quantitative speed of sound (SOS) data, which are translated into a higher risk when the scores are more negative. PRS-3 has an almost horizontal regression line due to the non-significant association.

The studied PRSs showed a moderate discriminative capacity, as evidenced by the areas under ROC curves (PRS-1 AUC, 0.645; PRS-2 AUC, 0.61; PRS-3 AUC, 0.526; PRS-4 AUC, 0.625) (Fig. 3).

DISCUSSION

In the present study, we tested four previously published PRSs derived from high-powered GWAS of various osteoporosis-related traits. By comparing the results in a group of patients with osteoporosis and in the control Spanish population, we demonstrate that three of the four PRSs tested are significantly associated with osteoporosis. So, our findings confirm that genetic profiling may help in the identification of osteoporosis, although their discriminative capacity is only moderate and their clinical relevance is still to be demonstrated.

The seminal study conducted by Estrada et al. back in 2012 was the largest GWAS on osteoporosis to that date and identified a total of 63 genetic variants associated with femur neck BMD. They showed that their genetic score predicted the risk of osteoporosis (1.56 odds for osteoporosis of women in the highest bin). The prediction ability was small, with an area under the curve (AUC) of 0.59 with the genetic score alone for osteoporosis (4). In the present study the odds ratio was 1.83 for osteoporosis in the last quartile with an AUC of 0.645 with the genetic score alone. Thus, these results appeared to be somewhat better compared to those from the original study. Hence, this predictive model might be of interest for the Spanish population.

More recently, Forgetta et al. created a PRS model trained with LASSO including 21 717 from a total of 345 111 SNPs significantly associated with ultrasound speed of sound (SOS), which decreased the number of people requiring CRF-FRAX and BMD-FRAX assessments (7). The prediction model, known as gSOS, can improve fracture risk prediction. Thus, a lower gSOS that is related to a lower SOS, was associated to a higher rate of major osteoporotic and hip fractures in European populations (8). The authors showed that the population in the quartile with the lowest gSOS had an odds ratio of 1.68 for major osteoporotic fracture risk and 1.57 for hip fracture. They also demonstrated that gSOS predicts major osteoporotic fracture and hip fracture with an AUC of 0.734 and 0.798,

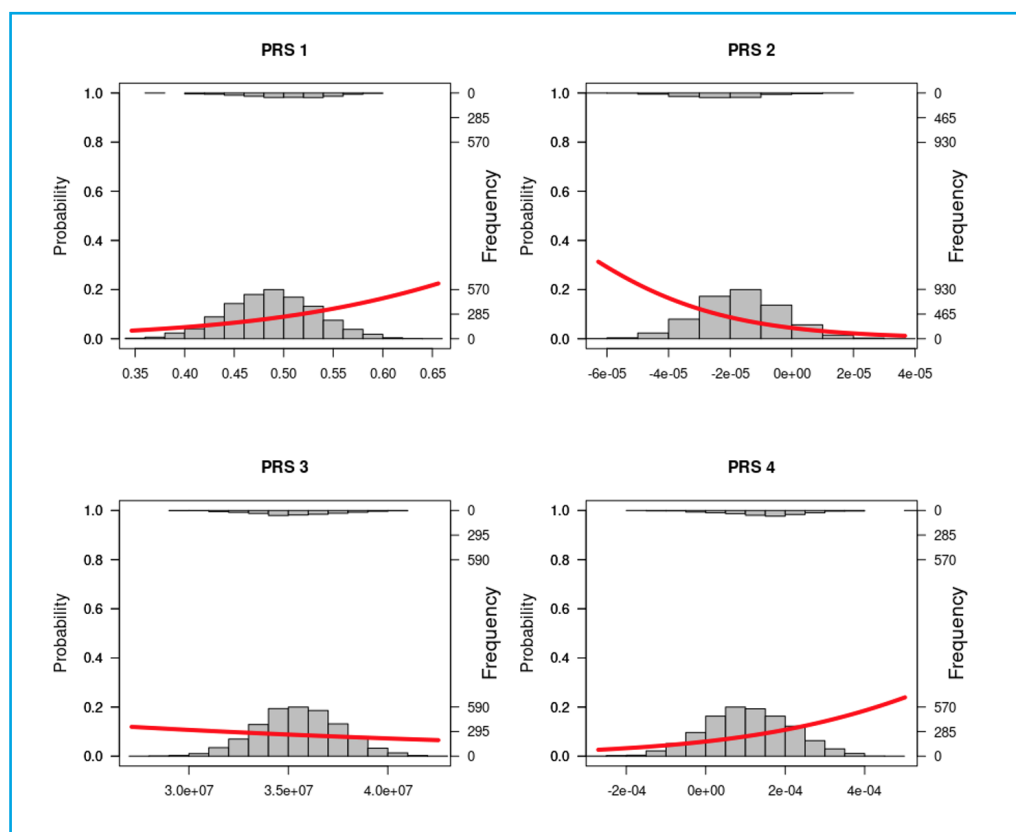


Figure 2. Frequency and linear logistic plots with the controls as zero (the bottom) and the cases as one (the top) for each PRS tested in the manuscript. Left axis is the probability of being control or cases according to the linear logistic regression. Whereas, the right axis shows the frequency of each group in a score range. The X axis scores corresponds to the PRS obtained.

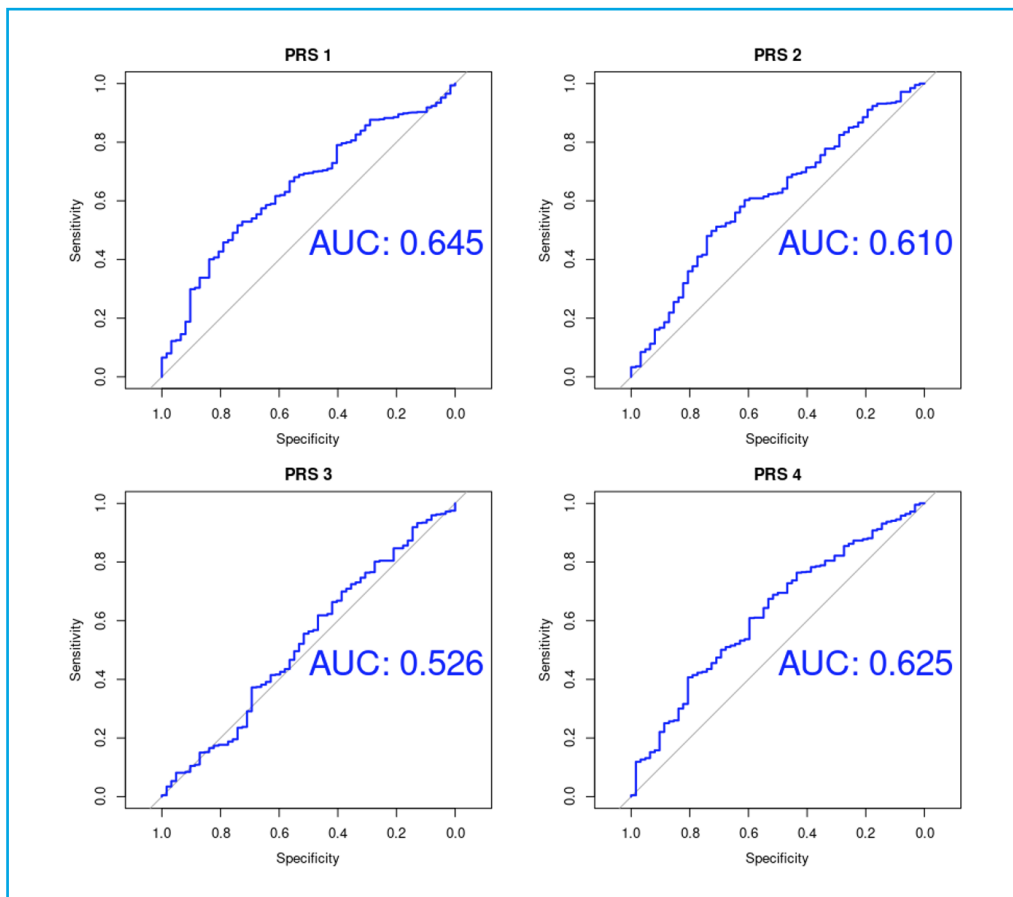


Figure 3. ROC curves for each PRS tested in the manuscript. Area under the curve (AUC) is shown in the center of each plot.

respectively. In our data, people with PRS score in the first quartile had a 2.11-fold risk of osteoporosis and an AUC of 0.61. The AUC is not as high as the one from the reference article. However, we could not use the fractures as a dependent variable because we did not have those data. Hence, our comparisons were not as clean as theirs in terms of the dependent variable.

The UK Biobank project is a noteworthy program that permits to study GWAS data with different phenotypes in a large number of individuals from the UK overall population (11). Thanks to these genetic and phenotypic data, various PRSs have been created. A study conducted by Rivas et al. proposed up to 813 PRSs models to predict over 1500 traits with genetic and phenotypic data from the UK Biobank including 2 PRSs models related to osteoporosis identified as PGS001273 and PGS001274 in the Polygenic Score Catalog (www.pgscatalog.org). Both have been trained with data from more than 260 000 individuals with European ancestry. The first one has been related to 'osteoporosis' ($n = 316$ SNPs) whereas the second one is associated with 'osteoporosis without pathological fracture' ($n = 1270$). Overall, they have found that the size of the PRS model is related to an increased predictive power. Thus, with the score model, they obtained an AUC of 0.629 and 0.718 in PGS001273 and PGS001274,

respectively (9). In our own study, PGS001273 did not show a significant association with osteoporosis. However, the largest PRS (PGS001274) was associated with osteoporosis (OR of the first quartile vs other quartiles 1.72), and a predictive power with an AUC of 0.625.

The heterogeneity between PRSs predictions is based on the heterogeneity between GWAS results, which are possibly different due to the use of distinct variables, outcome measurement, and the ancestry of samples (12). That is the reason why combining the results of the PRS tested does not improve the levels of prediction. For the same reasons, prediction models validated in each of these studies are somewhat better compared to ones obtained from our own data except for the study conducted by Estrada. This might also be associated with the fact that BMD was used as an outcome measure in Estrada's report and in the present study whereas calcaneal ultrasound was used in the UK Biobank-derived studies. Nevertheless, the predictive ability of the PRSs is low and, nowadays, they must be used as a complement for diagnosis with other clinical parameters such as FRAX scores.

This study has several limitations. We included a well-characterized group of individuals with primary osteoporosis of Spanish ancestry. However, our con-

control group obtained from the Spanish DNA biobank is well characterized according to the ancestry, sex and age. However, there are not data regarding clinical bone factors. As a matter of fact, some of them might present osteoporosis, which would decrease the study power. Also, the limited sample size, particularly of the patients' group, limit the statistical power of the study. This, and the lack of data about the controls, precluded the adjustment of the genetic associations by some relevant clinical factors.

In conclusion, several PRSs show significant differences between the overall Spanish population and patients with osteoporosis. This result supports the concept that PRSs may help identify individuals at risk of osteoporosis. Their exact role alone and in combination with other clinical factors remains to be elucidated.

REFERENCES

1. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 2018;50(9):1219-24. DOI: 10.1038/s41588-018-0183-z
2. Nguyen TV, Eisman JA. Post-GWAS Polygenic Risk Score: Utility and Challenges. *JBMR Plus* 2020;4(11):e10411. DOI: 10.1002/jbm4.10411
3. Uffelmann E, Huang QQ, Munung NS, de Vries J, Okada Y, Martin AR, et al. Genome-wide association studies. *Nat Rev Methods Primer* 2021;1(1):59. DOI: 10.1038/s43586-021-00056-9
4. Estrada K, Styrkarsdottir U, Evangelou E, Hsu YH, Duncan EL, Ntzani EE, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet* 2012;44(5):491-501. DOI: 10.1038/ng.2249
5. Morris JA, Kemp JP, Youlten SE, Laurent L, Logan JG, Chai RC, et al. An atlas of genetic influences on osteoporosis in humans and mice. *Nat Genet*. 2019;51(2):258-66. DOI: 10.1038/s41588-018-0302-x
6. Choi SW, Mak TS, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. *Nat Protoc*. 2020;15(9):2759-72. DOI: 10.1038/s41596-020-0353-1
7. Forgetta V, Keller-Baruch J, Forest M, Durand A, Bhatnagar S, Kemp JP, et al. Development of a polygenic risk score to improve screening for fracture risk: A genetic risk prediction study. *PLoS Med* 2020;17(7):e1003152. DOI: 10.1371/journal.pmed.1003152
8. Lu T, Forgetta V, Keller-Baruch J, Nethander M, Bennett D, Forest M, et al. Improved prediction of fracture risk leveraging a genome-wide polygenic risk score. *Genome Med* 2021;13(1):16. DOI: 10.1186/s13073-021-00838-6
9. Tanigawa Y, Qian J, Venkataraman G, Justesen JM, Li R, Tibshirani R, et al. Significant sparse polygenic risk scores across 813 traits in UK Biobank. *PLoS Genet* 2022;18(3):e1010105. DOI: 10.1371/journal.pgen.1010105
10. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;12(1):77. DOI: 10.1186/1471-2105-12-77
11. Collins R. What makes UK Biobank special? *Lancet Lond Engl* 2012;379(9822):1173-4. DOI: 10.1016/S0140-6736(12)60404-8
12. Yang TL, Shen H, Liu A, Dong SS, Zhang L, Deng FY, et al. A road map for understanding molecular and genetic determinants of osteoporosis. *Nat Rev Endocrinol* 2020;16(2):91-103. DOI: 10.1038/s41574-019-0282-7