



Article

# Using a Polygenic Score to Predict the Risk of Developing Primary Osteoporosis

Bulat Yalaev <sup>1,\*</sup>, Anton Tyurin <sup>2</sup>, Inga Prokopenko <sup>3</sup>, Aleksandra Karunas <sup>1</sup>, Elza Khusnutdinova <sup>1</sup> and Rita Khusainova <sup>1,4,5</sup>

- <sup>1</sup> Laboratory of Human Molecular Genetics, Institute of Biochemistry and Genetics-Subdivision of the Ufa Federal Research Centre of the Russian Academy of Sciences, 450054 Ufa, Russia  
<sup>2</sup> Internal Medicine Department, Bashkir State Medical University, 450008 Ufa, Russia  
<sup>3</sup> Section of Statistical Multi-Omics, Department of Clinical and Experimental Medicine, School of Biosciences and Medicine, University of Surrey, Guildford GU2 7XH, UK  
<sup>4</sup> Department of Genetics and Fundamental Medicine, Bashkir State University, 450008 Ufa, Russia  
<sup>5</sup> Department of Medical Genetics and Fundamental Medicine, Bashkir State Medical University, 450008 Ufa, Russia  
\* Correspondence: yalaev.bulat@gmail.com; Tel.: +7-347-235-60-88

**Abstract:** Osteoporosis (OP) is a multifactorial bone disease belonging to the metabolic osteopathies group. Using the polygenic score (PGS) approach, we combined the effects of bone mineral density (BMD) DNA loci, affecting osteoporosis pathogenesis, based on GEFOS/GENOMOS consortium GWAS meta-analysis. We developed models to predict the risk of low fractures in women from the Volga-Ural region of Russia with efficacy of 74% (AUC = 0.740; OR (95% CI) = 2.9 (2.353–3.536)), as well as the formation of low BMD with efficacy of 79% (AUC = 0.790; OR (95% CI) = 3.94 (2.993–5.337)). In addition, we propose a model that predicts fracture risk and low BMD in a comorbid condition with 85% accuracy (AUC = 0.850; OR (95% CI) = 6.6 (4.411–10.608)) in postmenopausal women.



**Citation:** Yalaev, B.; Tyurin, A.; Prokopenko, I.; Karunas, A.; Khusnutdinova, E.; Khusainova, R. Using a Polygenic Score to Predict the Risk of Developing Primary Osteoporosis. *Int. J. Mol. Sci.* **2022**, *23*, 10021. <https://doi.org/10.3390/ijms231710021>

Academic Editor: Alberto Falchetti

Received: 8 July 2022

Accepted: 30 August 2022

Published: 2 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** osteoporosis; polygenic risk score; fracture; low bone mineral density

## 1. Introduction

Osteoporosis is a systemic skeletal disease characterised by an increased risk of bone fractures, reduced bone mineral density (BMD) and disruption of the microarchitecture of cancellous and cortical bone [1]. The principal reason for the development of these processes is disruption of bone metabolism, in which catabolism and bone resorption begin to prevail over osteoblastogenesis, osteogenesis and bone mineralisation [2]. Primary osteoporosis includes postmenopausal, senile, juvenile and idiopathic types [3]. Secondary osteoporosis is associated with a number of comorbidities or with taking medication that can have a negative effect on bone remodelling [4]. It is estimated that 33% of women and 20% of men are at risk of developing an osteoporotic fracture, particularly in the hip, spine and forearm [5,6]. The WHO (World Health Organization) operational definition of osteoporosis is based on a DXA measurement of bone mineral density, based on the clear link between lower BMD and increased fracture risk [7,8].

Three groups of drugs are used to treat OP [9,10]: stimulants of bone matrix formation, inhibitors of bone resorption, dual-action drugs. However, despite effective therapy for OP, early diagnosis of the disease remains a concern. There has been no reduction in fracture frequency and the number of patients with diagnosed osteoporosis is increasing worldwide, despite the availability of Garvan, FRAX and QFracture 10-year probability calculators [11–13]. In replicative studies, the AUC (Area Under Curve) of the Garvan and FRAX models ranged from 0.61 to 0.85, with a mean value of ~0.7. One would want the predictive power to be greater than 80%. Diagnostic systems for osteoporosis with better predictive potential need to be developed, e.g., using genetic studies [11].

The GEFOS/GENOMOS consortium on genome-wide search for genetic markers of primary osteoporosis identified 56 loci associated with low bone mineral density, 14 of which were associated with the risk of osteoporotic fractures [8]. These results resulted in an unexpected conclusion: fractures and low BMD may be independent endophenotypes of OP, which changes the fundamental ideas regarding the aetiology of the disease and requires new approaches to the diagnosis and therapy of OP [14]. However, the results of GWAS replication within this consortium, as well as in a number of other independent studies have not reproduced many significant associations [15]. This is particularly well demonstrated by the data for a sample of women from the Volga-Ural region of Russia (VUR) [8]. In the summary data of genetic effects and their confidence intervals, the most significant DNA variants from the population of European and Oriental origin were not the main risk contributors for a sample of women with OP from the Volga-Ural region [8,15]. VUR is an ethnically diverse region on the border between Europe and Asia near Russia's Ural Mountains. That is home to Finno-Ugric, Turkic and Slavic peoples, who have European and Mongoloid components in their gene pool in various proportions. The results of population research in this region agree with the data demonstrating the largest European contribution to the gene pool of the Mordvins, Komi and Russians and an increased Mongoloid component in the Turkic populations of the VUR [16].

This indicates the heterogeneous genetic structure of osteoporosis and the superior contribution of the ethnic component to the molecular pathogenesis of the disease. This actualises the definition of the DNA markers with the largest effect on osteoporosis and recognises that the implementation of advanced methods of complex statistical analyses is even more significant when studying multifactorial diseases in populations with a unique gene pool and ethnic structure, as compared to Western European populations [14]. Thus, there is a specific knowledge gap in both understanding the contribution of genetic factors to the risk of osteoporosis in ethnically differentiated populations and in understanding how the findings from GWAS and gene-candidate studies can be applied to improve the prognostic power of the current early osteoporosis diagnosis system.

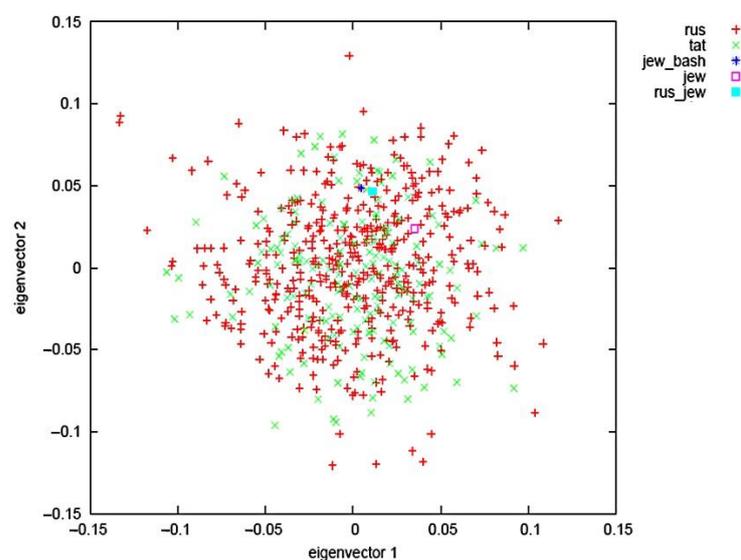
One method employed to assess the contribution of a large number of loci to disease risk is the polygenic score (PGS) approach. In this method, the effects of the associated variant risk alleles (OR/beta) are combined into a PGS, reflecting part of an individual susceptibility to disease in a population under study. More specifically, the PGS is based on the sum of the effects of independent risk variants associated with the disease, using current data from the largest and most informative genome-wide association studies [17,18]. Most studies in this direction are related to cardiovascular diseases, type 2 diabetes, breast and prostate cancer, as well as Alzheimer's disease [19]. Simultaneously, there are only a few studies dedicated to the polygenic risk definition of osteoporosis [20,21]. In clinical practice, there is a growing need to implement PGS for early diagnosis of OP prior to the occurrence of fractures [8]. The polygenic risk method was first successfully applied in the GWAS study of schizophrenia. In this study, PGS explained approximately 3% of the variance in schizophrenia [22]. Currently, this method is used to predict the biological properties of crops [23] and livestock [24], and to assess the risk of disease in humans [17]. Currently, the most convincing evidence of the effectiveness of PGS in response to treatment is the use of statins to reduce the risk of a first coronary event, as studies have confirmed that the relative risk reduction is higher in individuals with a high genetic risk of cardiovascular disease [25,26]. Large samples of unrelated individuals in human populations and random gene drift allow this method of analysis to be employed more efficiently; however, it also requires careful quality control of the data studied [23,27]. The use of PGS for the development of prognostic models of fractures and low BMD separately and combined, as well as the search for the most sensitive and specific risk models for primary osteoporosis appear to be relevant.

The aim of this study is to assess the polygenic risk of fractures and low BMD based on the analysis of 140 polymorphic loci of microRNA target sites, polymorphic variants of

candidate genes, along with genes associated with primary osteoporosis. We focused our research on the GWAS results for women with postmenopausal osteoporosis.

## 2. Results

To determine whether we would need to account for the population substructure within the studied sample of women, who predominantly self-reported as being of Tatar and Russian descent, we conducted principal component analysis (PCA), by using the information from the independent loci studied. Individuals with a 97% success rate of genotyping and SNPs with a 97% success rate of genotyping were included in the analysis. In general, no population stratification of the studied sample by the studied loci depending on their ethnicity was observed (Figure 1).

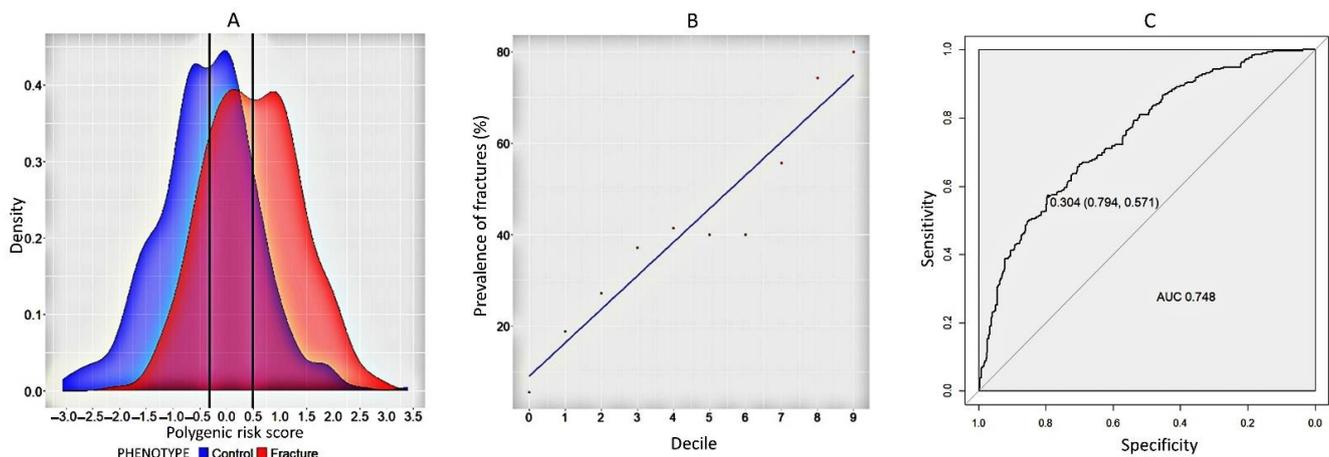


**Figure 1.** Location of women from the VUR for two principal components defined by Eigen vectors and the ethnic origin.

Overall, the sample of women was not genetically heterogeneous, which supported our analysis of polygenic risk assessment of fractures and low BMD without dividing into subgroups by ethnic origin, i.e., the analysis was performed as for a single population. The QC, involving the exclusion of loci that were in linkage disequilibrium and do not support the Hardy-Weinberg equilibrium, resulted in 140 loci out of a total of 150 loci being tested. The range of minor allele frequencies of these loci ranged from 1% to 49%, which was consistent with the required quality control requirements for PGS.

### 2.1. Model for Predicting Fracture Risk in Women

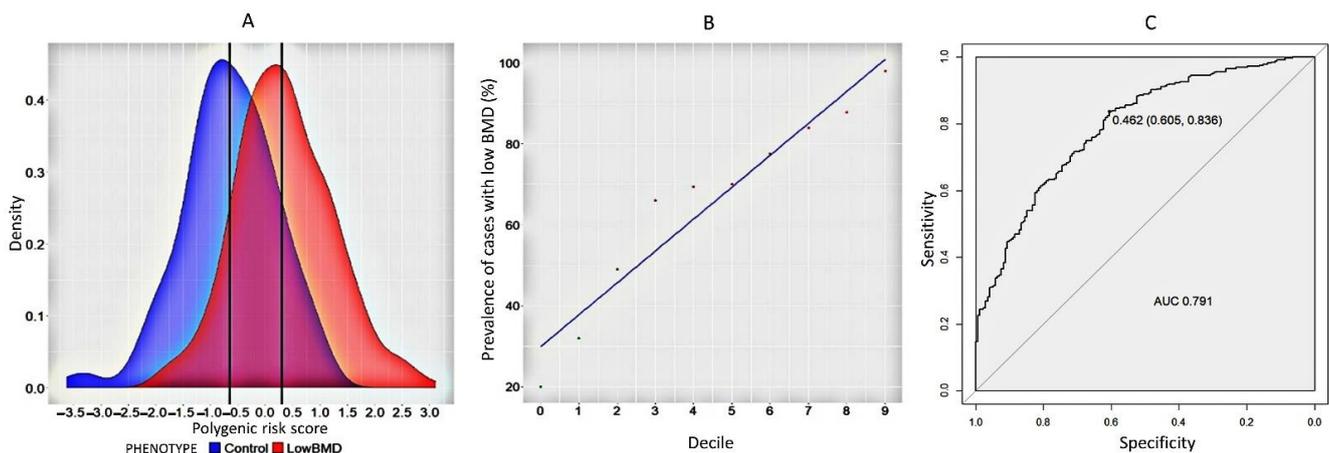
We calculated the PGS for each individual and assessed the relationship between PGS and osteoporotic fracture. We plotted the density distribution of indicators between the control and experimental groups (Figure 2A). The overwhelming proportion of polygenic indicators in women with fractures (median = 0.494) exceeded those in women without fractures (median = 0.357). The decile plot dividing the polygenic scale into 10% groups reveals that the polygenic risk index increases linearly with increasing fracture frequency (Figure 2B). ROC analysis demonstrated a high AUC (95% CI = 0.75 (0.71–0.78), sensitivity = 0.79 and specificity = 0.57 of the model (Figure 2C). Association analysis was statistically significant and indicated that the probability of fracture in the experimental group was 2.9 times higher (95% CI = 2.35–3.53),  $p$ -value =  $2.37 \times 10^{-16}$ ) compared to the control group.



**Figure 2.** (A) The graph of the density distribution of polygenic risk score in the control and cases in general; (B) decile plots by PGS within each decile for cases (fracture) and controls in women in the general sample; (C) ROC-curve relating to the PGS of fractures in women participating in this study.

### 2.2. Model to Predict the Formation of Low BMD in Women

According to the GEFOS/GENOMOS full genome study, it was determined that low BMD can be a separate endophenotype of osteoporosis, independent of osteoporotic fractures [8]. Therefore, it is of interest to study the contribution of the polygenic risk of this phenotype to OP. The analysis proved that the probability of forming a low BMD level was 3.94 times higher in the group with low mineral density compared to controls (95% CI = 2.99–5.34). To assess the difference in the distribution of the polygenic scores between the case and control groups, a plot of the density distribution of the scores between the control and experimental groups was plotted using R software (Boston, MA, US) (Figure 3A). As can be noted from the graph, the overwhelming proportion of polygenic scores in women with low BMD (mean value = 0.35) exceeded those in women with normal BMD (mean value = −0.65). In addition, we plotted the deciles of polygenic scores as a function of fracture incidence (Figure 3B). We report relatively high AUC = 0.791, sensitivity = 0.605 and specificity = 0.836 of the model ( $p$ -value =  $2.54 \times 10^{-16}$ , Figure 3C).

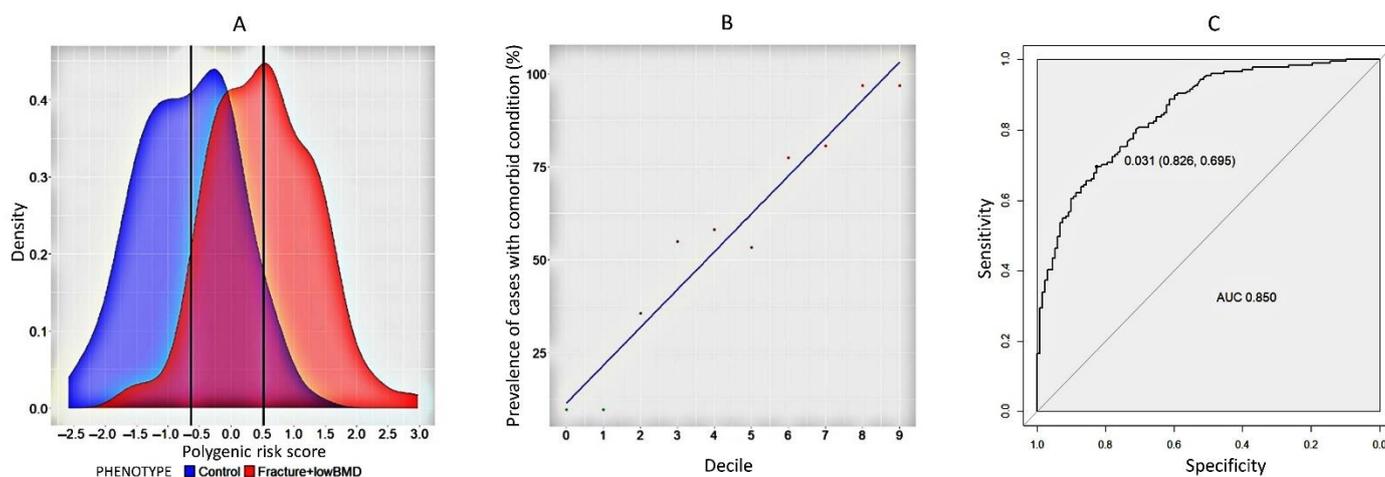


**Figure 3.** (A) The graph of the distribution of the density of polygenic risk score for the formation of a low level of BMD in the control (blue) and cases (red) in general; (B) decile plots by PGS within each decile for cases (low BMD) and controls in women in the general. (C) ROC-curve analysis of the polygenic risk assessment of the formation of a low level of BMD in women in the general sample.

### 2.3. A Model to Predict Fractures and Low BMD in Comorbid Women

Osteoporosis is a combination of two phenotypes—fractures and low BMD. Therefore, we developed a prognostic polygenic risk model in which women with both traits in a

comorbid condition were analysed. At this stage, the study sample included patients with per-fractures and low BMD, therefore, a separate analysis of these cases was performed to determine the risk of these combined conditions. As can be established from the density distribution plot of the polygenic risk scale, the polygenic score of women with low BMD and fractures (median value 0.511) in the combined condition is higher than that of women with normal BMD and no fractures (median value  $-0.686$ ) (Figure 4A). The decile plot illustrates a linear increase in the incidence of cases with the combined phenotype (Figure 4B). Sensitivity and specificity analysis of the model shows a high AUC (95% CI = 0.850 0.809–0.892), sensitivity = 0.826 and specificity = 0.695 (Figure 4C). Association analysis of the model confirmed a 6.6-fold (95% CI: 4.411–10.608) higher probability of fracture compared with the control group ( $p = 2.76 \times 10^{-16}$ ). The model has the highest prognostic potential in the total sample of women.



**Figure 4.** (A) The graph of the distribution of the density of PGSs of fracture with a low level of BMD in the control group and cases in women; (B) decile plots by PGS within each decile for cases and controls in women in the general; (C) ROC-curve analysis of PGS of fractures with a low level of BMD in women in the general sample.

### 3. Discussion

Based on the results obtained in this work together with the GWAS replication data within the GEFOS consortium, we first attempted to develop PGS models that predict low BMD and risk of osteoporotic fractures separately and in a comorbid condition in women from the VUR. Overall, our study findings are in each resulting model, the top percentile accounted for people at high risk of low-traumatic fractures and low BMD. It was determined that all the constructed models have high prognostic significance, allowing for a high degree of probability to correctly classify patients and the control group, as well as correctly, based on numerical indicators, differentiate differences on a polygenic risk scale. The highest indicators of sensitivity and specificity were exhibited by models of fracture risk and low BMD in combined conditions (AUC > 0.800).

Comparison with previous studies and the research value of our results. These results are new and interesting, given that the predictive models obtained in previous studies had less predictive value. For example, the use of the Lasso algorithm to construct a genetic risk scale in Li et al. (2019), established that the AUC area reduced from 0.653 to 0.587 and 0.588 in models that included clinical factors contributing to lower femoral neck BMD [28]. The authors emphasised the limited effectiveness of current genetic studies that focus on people of European descent when trying to improve the prediction of multifactorial disease and called for increasingly extensive studies targeting other ethnic groups [28]. Using data from the British Biobank, Tianyuan Lu et al. (2021), developed a polygenic risk scale of 21,717 genetic variants ( $r^2 = 23.2\%$ ) explaining 25% of the total fracture risk variance [20]. This significantly exceeded the previous estimate of 22,886 SNPs, which explained 17.2% of

the total variance and achieved an AUC value of 0.570 when studying patients with any type of fracture [21]. In comparison, the clinical risk cofactors FRAX, age, sex and their combination explained only 4% of the polygenic risk variance. PGS has been shown to be independent of other commonly used clinical risk factors and to be significantly associated with fracture risk [29]. However, it is recognised that the polygenic risk markers relevant to populations of European origin are not quite effective for prognostic models of fractures and the formation of low BMD in Chinese populations.

According to our data, low-traumatic fractures occur on the upper percentile, which is consistent with the results obtained by Tian-Yuan Lu et al. (2021), who explained that individuals with common fractures belonged to the group below the third percentile (2.5%) (95%  $\approx$  1.5–3.6%). Simultaneously, the incidence of severe osteoporotic fractures and hip fractures was only 0.7% (95% CI 0.2–1.4%) and 0.3% (95% CI 0–0.7%), respectively, among individuals with a PGS above the 99th percentile. Thus, polygenic risk assessment effectively predicts cumulative fracture risk, with low-traumatic fractures characterised by an increased polygenic risk [20]. Kim (2018) analysed a cohort from the British Biobank, in which a PGS of low BMD formation was performed based on LASSO analysis. There was a tendency for the risk of osteoporosis to decrease dramatically with increasing LASSO scores. For example, those with a low LASSO score of 4 (2.2% of individuals) had a 19-fold higher risk of developing OP than the middle cell (95% CI (6.18–6.20)). Those with the highest risk of OP were 0.221 (95% CI (0.218–0.224)) relative to the median [21].

The results of the HoLe et al. (2016) study, which genetically assessed the risk of per-relapse fractures and decreased BMD of the femoral neck and lumbar spine, are of interest. Based on data on 62 polymorphic variants, the authors revealed that the probability of fracture was significantly higher in men and women when the polygenic risk value was  $>4.24$ . In the Cox Proportional Hazards Model, an increase in polygenic risk of decreased femoral neck BMD was associated with a 20% increase in fracture risk. In addition, analysis by fracture type showed that the relationship between femoral neck polygenic risk and fracture risk was observed for spine and wrist fractures, but not for femoral neck fractures. The magnitude of the relationship remained almost unchanged after adjustment for age, sex, previous fracture and number of falls in the previous 12 months. However, after an additional correction of the BMD level of the femoral neck, the association between the polygenic risk of decreased BMD and fractures in general was no longer statistically significant [30]. The small effect size of polymorphic variants is to be expected according to a simple epidemiological principle [31]. Research data suggest that with an OR of up to 1.2 and assuming a 5-year per-relapse rate of 10%, the area under the ROC curve associated with this genetic variant is expected to be about 0.55. Thus, a single risk allele alone would not be helpful in predicting the risk of OP. At the same time, a combination of several genetic variants may be useful even for modest effect sizes, because the discriminatory power of the AUC increases in proportion to the effect size [32]. For example, when summing the effects of 50 genetic variants with an OR = 1.1 size, the AUC would be roughly 0.63; with an OR = 1.2, the AUC is expected to be about 0.73. It is assumed that a model with 500 variants, each with an OR = 1.1, could result in an AUC of up to 0.80.

Polygenic risk assessment is designed to improve the quality of predictive medicine [33], although its results should be interpreted with extreme caution, taking into account the ethnic origin of the sample and the variety of bioinformatic approaches to statistical analysis [34]. Thus, the results of PGS can be valuable in diagnostic medicine and have higher prognostic power at the individual level. Nonetheless, previously obtained results have low prognostic power, which generates considerable debate as regards their direct clinical benefit [35]. Consequently, it is necessary to continue research in this way, as the clinical implementation of prognostic methods for osteoporosis can be beneficial in cases where there is a higher a priori probability of disease, for example, in the early stages of the disease, to help in the diagnosis or to inform the choice of treatment [17].

Clinical value of the data obtained. The development of prognostically accurate models of osteoporosis genetics depends less on the coverage of the number of loci studied

and the power of statistical research parameters but rather on studying the relationship between loci and OP phenotypes in individual bone localisations, which allows us to take a narrower approach to the tactics of osteoporosis research. In terms of finding risk markers, it is interesting to study the weighted effect of many markers on an individual's risk of developing the disease, which can be measured using polygenic risk assessment. The efficiency of osteoporosis prognosis using polygenic assessment has increased. This is probably due to the absence of measurement error inherent in the conventional risk calculation using FRAX. In addition, PGS is independent of other clinical FRAX risk factors, and finally, genotypes are measured with very high accuracy. It is currently undetermined which bioinformatic analysis methods are best suited for predicting osteoporosis. There is a high contribution of the ethnic component to the genetic factor in the development of osteoporosis; in addition, indicators can be influenced by different proportions of heritable ethnic components, which can lead to a lower proportion of sensitivity models in a sample of women. Since allele and genotype frequencies can vary between ethnic groups, each allele can be at one frequency in one ethnic group but less frequent in another. Consequently, the effect size (e.g., odds ratio) for a risk allele may differ for different ethnic groups, particularly given different environmental factor backgrounds. Population characteristics (such as size, founder influence and inbreeding), as well as inheritance patterns may also differ. Thus, a polygenic risk assessment constructed when studying one ethnic group may not be applicable to another. To date, most PGS osteoporosis studies have been conducted in populations of European origin. However, each of the models studied in this research has a high prognostic value and can be used in the future for replicative studies and the development of promising early diagnostic systems based on the individual genetic profile. Nevertheless, the observed relationship between genotypes and continuous traits was largely linear, making methods based on polygenic risk assessment and regression analysis potential avenues for early diagnosis.

A limitation of the study. A limitation related to our study is the small sample size, the number of SNPS studied and the lack of inclusion of clinical predictors in the model studied. In addition, the results have not been tested on other cohorts from the Volga-Ural region. However, the main goal of the study was to develop a polygenic risk scale specifically for the study population of women from the Volga-Ural region who participated in the full genome-wide GWAS study and to understand the potential applicability of this method in the future for early diagnosis of osteoporosis. In the future, we plan to expand the sample with a polygenic risk assessment and to perform this analysis on other cohorts of patients with postmenopausal osteoporosis.

## 4. Methods and Materials

### 4.1. Phenotypic Information

The case-control study involved 701 postmenopausal women (mean age =  $61.95 \pm 7.94$ ), who underwent medical examination between 2004 and 2011 and lived in the regions of Ekaterinburg and Ufa, Russia. There were women with primary osteoporosis and age-matched control groups without fractures and with a normal level of BMD. The study included 294 women with fractures and 407 without fractures. BMD level was measured in 496 women, including the femoral neck in 270 women and the lumbar spine in 281 women. We completely excluded the presence of any family groups and relatives in the sample on the basis of questionnaires and family history data. The clinical characteristics of the groups are presented in Table 1. Exclusion criteria comprised current use of bisphosphonates or other medications known to affect bone metabolism, a history of alcohol and drug abuse, chronic glucocorticoid use, hormone replacement therapy, current treatment for acute medical condition, bilateral knee and/or hip replacement, metal hardware in both lower extremities, history of heterotopic ossification at the hip and knee, history of cardiovascular disease, besides a history of diseases known to affect bone metabolism (e.g., ankylosing spondylitis, rheumatoid arthritis, Paget's disease, hyperparathyroidism, hypogonadism, hypothyroidism, diabetes mellitus and/or cancer). The BMD level was measured by

biphasic absorption X-ray densitometry (DEXA) using a Hologic QDR 4500/A DXA system (USA) in standard localisations (femoral neck and lumbar spine). The general sample was divided according to T-criterion—a T-value of more than  $-1.0$  described a normal BMD, values from  $-1.0$  to  $-2.5$  diagnosed osteopenia, whilst values less than  $-2.5$  diagnosed osteoporosis (in accordance with the WHO recommendations) [1], and also according to the presence or absence of osteoporetic fractures in standard localisations (axial femur, lumbar spine). Each participant signed an informed consent form to participate in the study, in accordance with the standards developed by the World Medical Association Declaration of Helsinki “Ethical Principles for Scientific Medical Research Involving Human Subjects” and with the approval of the local Bioethics Committee of the Institute of Biochemistry and Genetics UFRC RAS.

**Table 1.** Characteristics of the study by phenotype subgroups.

Parameters	Woman		
	N	Age, Years, Me $\pm$ SD	BMI, kg/m <sup>2</sup> , Me $\pm$ SD
With fractures	294	62.16 $\pm$ 7.95	27.00 $\pm$ 3.60
Without fractures	407	60.14 $\pm$ 8.01	27.80 $\pm$ 3.81
With low BMD	324	62.17 $\pm$ 7.95	27.04 $\pm$ 3.60
With normal BMD	172	60.23 $\pm$ 7.97	28.10 $\pm$ 4.90
With low BMD and with fractures	177	60.15 $\pm$ 7.43	27.00 $\pm$ 2.40
With normal BMD and without fracture	260	61.76 $\pm$ 7.96	28.90 $\pm$ 3.70

Abbreviations: Me—mean value, SD—standard deviation.

#### 4.2. Genetic Data, Quality Control and Analysis

The study included 140 polymorphic variants and GWAS replication results within GEFOS (Genetic Factors in Osteoporosis Consortium) [8] and was based on the findings (genotypes and their effects) from previous genetic studies of osteoporosis in the VUR (Supplementary Data), including polymorphic variants of structural genes, transcription factors, steroid hormone receptors, signalling pathways, enzymes, intergenic sites, microRNA genes and their binding sites in addition to other functional regions (Tables S1–S4). Genotyping of the studied loci was conducted using polymerase chain reaction (PCR) “in real time” by way of using TaqMan technology on the platform of the CFX96 device (Biorad Touch Real-Time PCR Detection System).

#### 4.3. Statistical Analysis

Genetic heterogeneity of the studied sample was assessed using principal component analysis (PC) in the Smartpca software [36]. Heterogeneity was assessed using the Cochran Q-criterion, differences were considered statistically significant at  $p < 0.1$ , heterogeneity level was revealed using  $I^2$  criterion and filtering was performed using PLINK 1.09 software (Cambridge, MA, USA) [37]. Linkage disequilibrium analysis and Hardy-Weinberg equilibrium compliance ( $p < 0.05$ ) were completed using Haploview 4.2 software (Cambridge, MA, USA) [38]. PGS was undertaken using PLINK 1.09 software and was calculated by summing up the effects of risk alleles in everyone, weighted by the effect size of the risk alleles (OR). PGSs obtained for each individual were further normalised using the normalisation script of the R-studio software version 1.4.1717 (Boston, MA, USA).

The formula for calculating PGS is as follows [39]:

$$PRS = \frac{\sum_i^N S_i * G_{ij}}{P * M_j}$$

where  $S_j$  is the effect size (OR);  $G_{ij}$  is the observed number of risk alleles;  $P$  is the sampling density (2 for humans);  $N$  is the total number of polymorphic variants;  $M_j$  is the number of missing genotypes in the analysed array.

The PGS assessment included distinctive five steps:

1. Quality control of baseline data with the exclusion of polymorphic loci that are linkage disequilibrium and do not support the Hardy-Weinberg equilibrium.
2. Analysis of polygenic risk assessment.
3. PGS specificity and sensitivity analysis using ROC-analysis.
4. The search for the most sensitive and specific model of polygenic fracture risk and low BMD was completed on the basis of ROC-analysis on the basis of RStudio, in which the sensitivity and specificity of the model as well as the AUC, a numerical index of the area under the ROC-curve, were assessed.

## 5. Conclusions

In conclusion, we generated a polygenic risk score for primary osteoporosis that is consistently associated with fracture risk and low BMD in postmenopausal women from the Volga-Ural region of Russia. By analysing the 3 prognostic models, we found that the model including women with fractures combined with low BMD had the highest sensitivity and specificity. The loci included in the models for analysis are of interest for further research. PGS should be studied as a tool to provide better risk stratification to identify individuals at high risk of fracture, especially given that polygenic risk estimates can be derived from effective early diagnostic methods.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms231710021/s1>.

**Author Contributions:** Conceptualization, B.Y. and R.K.; methodology, I.P.; software, B.Y.; validation, B.Y.; formal analysis, B.Y., A.T., R.K.; investigation, B.Y.; resources, R.K.; data curation, R.K., I.P.; writing—original draft preparation, B.Y.; writing—review and editing, B.Y., R.K., A.T.; visualization, B.Y.; supervision, R.K.; project administration, A.K., E.K.; funding acquisition, A.K., E.K.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was carried out within the framework of the state assignment of the Ministry of Science and Higher Education of the Russian Federation (No. 122041400169-2), with partial support from the Megagrant of the Government of the Russian Federation (Agreement no. 075-15-2021-595) and from the grant in the field of science from the budget of the Republic of Bashkortostan for state support of young scientists, post-graduate students and candidates of science (competition code—NOC-GMU-2021). DNA samples for the study have been taken from the Collection of Human Biological Materials of the IBG UFRC RAS supported by the Program of Bioresource Collections of the Russian Federal Agency of Scientific Organizations (Agreement No. 007-030164/2) using the equipment of Regional center for collective use «Agidel».

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of the Institute of Biochemistry and Genetics (protocol code: 14; date: 15 September 2016).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The additional data are available in Supplementary.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Kanis, J.; Cooper, C.; Rizzoli, R.; Reginster, J. European guidance for the diagnosis and management of osteoporosis in postmenopausal women. *Osteoporos. Int.* **2019**, *30*, 3–44. [[CrossRef](#)] [[PubMed](#)]
2. Morris, J.A.; John, P.K.; Scott, E.Y.; Laetitia, L.; John, G.L.; Ryan, C.C.; Nicholas, A.V.; Forgetta, V.; Kleinman, A.; Mohanty, S.T.; et al. An atlas of genetic influences on osteoporosis in humans and mice. *Nat. Genet.* **2019**, *51*, 258–266. [[CrossRef](#)] [[PubMed](#)]
3. Dobbs, M.B.; Buckwalter, J.; Saltzman, C. Osteoporosis: The increasing role of the orthopaedist. *Iowa Orthop. J.* **1999**, *19*, 43–52. [[PubMed](#)]
4. Panday, K.; Gona, A.; Humphrey, M.B. Medication-induced osteoporosis: Screening and treatment strategies. *Ther. Adv. Musculoskelet Dis.* **2014**, *6*, 185–202. [[CrossRef](#)]

5. Hanxuan, L.; Zhousheng, X.; Quarles, L.D.; Wei, L. Osteoporosis: Mechanism, molecular target and current status on drug development. *Curr. Med. Chem.* **2020**, *28*, 1489–1507. [[CrossRef](#)]
6. Wiedl, A.; Stefan, F.; Annabel, F.; Edgar, M. Fractures' associated mortality risk in orthogeriatric inpatients: A Prospective 2-year survey. *Eur. Geriatr. Med.* **2021**, *1*, 61–68. [[CrossRef](#)]
7. Marshall, D.; Johnell, O.; Wedel, H. Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. *Br. Med. J.* **1996**, *312*, 1254–1259. [[CrossRef](#)]
8. Estrada, K.; Stykarsdottir, U.; Evangelou, E.; Hsu, Y.H.; Duncan, E.L.; Ntzani, E.E.; Oei, L.; Albagha, O.M.; Amin, N.; Kemp, J.P.; 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*, 491–501. [[CrossRef](#)]
9. Khosla, S.; Hofbauer, L.C. Osteoporosis treatment: Recent developments and ongoing challenges. *Lancet Diabetes Endocrinol.* **2017**, *5*, 898–907. [[CrossRef](#)]
10. Marie, P.J. Strontium ranelate: A dual mode of action rebalancing bone turnover in favour of bone formation. *Curr. Opin. Rheumatol.* **2006**, *18*, 11–15. [[CrossRef](#)]
11. Bolland, M.J.; Siu, A.T.; Mason, B.H.; Horne, A.M.; Ames, R.W.; Grey, A.B.; Gamble, G.D.; Reid, I.R. Evaluation of the FRAX and Garvan fracture risk calculators in older women. *J. Bone Miner. Res.* **2011**, *26*, 420–427. [[CrossRef](#)] [[PubMed](#)]
12. Sandhu, S.K.; Nguyen, N.D.; Center, J.R.; Pocock, N.A.; Eisman, J.A.; Nguyen, T.V. Prognosis of fracture: Evaluation of predictive accuracy of the FRAX algorithm and Garvan nomogram. *Osteoporos. Int.* **2010**, *21*, 863–871. [[CrossRef](#)] [[PubMed](#)]
13. Hippisley-Cox, J.; Coupland, C. Derivation and validation of updated QFracture algorithm to predict risk of osteoporotic fracture in primary care in the United Kingdom: Prospective open cohort study. *BMJ* **2012**, *344*, 3427. [[CrossRef](#)] [[PubMed](#)]
14. Yang, T.-L.; Shen, H.; Liu, A.; Dong, S.-S.; Zhang, L.; Deng, F.-Y.; Zhao, Q.; Deng, H.-W. A road map for understanding molecular and genetic determinants of osteoporosis. *Nat. Rev. Endocrinol.* **2020**, *16*, 91–103. [[CrossRef](#)] [[PubMed](#)]
15. Liu, Y.-J.; Zhang, L.; Papiasian, C.J.; Deng, H.-W. Genome-wide association studies for osteoporosis: A 2013 update. *J. Bone Metab.* **2014**, *21*, 99–116. [[CrossRef](#)]
16. Khusnutdinova, E.K.; Victorova, T.V.; Fatkhislamova, R.I.; Galeeva, A.R. Relative contributions of the caucasoid and mongoloid components to the formation of ethnic groups of the Volga-Ural region as estimated from DNA polymorphism. *Genetika.* **1999**, *35*, 1132–1137.
17. Lewis, C.M.; Evangelos, V. Polygenic risk scores: From research tools to clinical instruments. *Genome Med.* **2020**, *12*, 44. [[CrossRef](#)]
18. Torkamani, A.; Wineinger, N.E.; Topol, E.J. The personal and clinical utility of polygenic risk scores. *Nat. Rev. Genet.* **2018**, *19*, 581–590. [[CrossRef](#)]
19. Lambert, S.A.; Abraham, G.; Inouye, M. Towards clinical utility of polygenic risk scores. *Hum. Mol. Genet.* **2019**, *28*, R133–R142. [[CrossRef](#)]
20. Lu, T.; Forgetta, V.; Keller-Baruch, J.; Nethander, M.; Bennett, D.; Forest, M.; Bhatnagar, S.; Walters, R.G.; Lin, K.; Chen, Z.; et al. Improved prediction of fracture risk leveraging a genome-wide polygenic risk score. *Genome Med.* **2021**, *13*, 1–15. [[CrossRef](#)]
21. Kim, S.K. Identification of 613 new loci associated with heel bone mineral density and a polygenic risk score for bone mineral density, osteoporosis and fracture. *PLoS ONE* **2018**, *13*, e0200785. [[CrossRef](#)] [[PubMed](#)]
22. International Schizophrenia Consortium; Purcell, S.M.; Wray, N.R.; Stone, J.L.; Visscher, P.M.; O'Donovan, M.C.; Sullivan, P.F.; Sklar, P. Common polygenic variation contributes to risk of schizophrenia that overlaps with bipolar disorder. *Nature* **2009**, *460*, 748. [[CrossRef](#)] [[PubMed](#)]
23. Heslot, N.; Jannink, J.L.; Sorrells, M.E. Perspectives for genomic selection applications and research in plants. *Crop Sci.* **2015**, *55*, 1–12. [[CrossRef](#)]
24. Hayr, M.K.; Druet, T.; Garrick, D.J. 027 Performance of genomic prediction using haplotypes in New Zealand dairy cattle. *J. Anim. Sci.* **2016**, *94*, 13. [[CrossRef](#)]
25. Natarajan, P.; Young, R.; Stitzel, N.O.; Padmanabhan, S.; Baber, U.; Mehran, R.; Sartori, S.; Fuster, V.; Reilly, D.F.; Butterworth, A.; et al. Polygenic risk score identifies subgroup with higher burden of atherosclerosis and greater relative benefit from statin therapy in the primary prevention setting. *Circulation* **2017**, *135*, 2091–2101. [[CrossRef](#)]
26. Schunkert, H.; Samani, N.J. Statin treatment: Can genetics sharpen the focus? *Lancet* **2015**, *385*, 2227–2229. [[CrossRef](#)]
27. Duncan, L.; Shen, H.; Gelaye, B.; Meijisen, J.; Ressler, K.; Feldman, M.; Peterson, R.; Domingue, B. Analysis of polygenic risk score usage and performance in diverse human populations. *Nat. Commun.* **2019**, *10*, 3328. [[CrossRef](#)]
28. Li, Y.M.; Peng, C.; Zhang, J.G.; Zhu, W.; Xu, C.; Lin, Y.; Fu, X.Y.; Tian, Q.; Zhang, L.; Xiang, Y.; et al. Genetic risk factors identified in populations of European descent do not improve the prediction of osteoporotic fracture and bone mineral density in Chinese populations. *Sci. Rep.* **2019**, *9*, 6086. [[CrossRef](#)]
29. Forgetta, V.; Keller-Baruch, J.; Forest, M.; Durand, A.; Bhatnagar, S.; Kemp, J.P.; Nethander, M.; Evans, D.; Morris, J.A.; Kiel, D.P.; et al. Development of a polygenic risk score to improve screening for fracture risk: A genetic risk prediction study. *PLoS Med.* **2020**, *17*, e1003152. [[CrossRef](#)]
30. Ho-Le, T.P.; Center, J.R.; Eisman, J.A.; Nguyen, H.T.; Nguyen, T.V. Prediction of bone mineral density and fragility fracture by genetic profiling. *J. Bone Miner. Res.* **2017**, *32*, 285–293. [[CrossRef](#)]
31. Holtzman, N.A.; Marteau, T.M. Will genetics revolutionize medicine? *N. Engl. J. Med.* **2000**, *343*, 141–144. [[CrossRef](#)] [[PubMed](#)]
32. Pepe, M.S.; Gu, J.W.; Morris, D.E. The potential of genes and other markers to inform about risk. *Cancer Epidemiol. Biomarkers Prev.* **2010**, *19*, 655. [[CrossRef](#)]

33. Sugrue, L.P.; Desikan, R.S. What are polygenic scores and why are they important? *JAMA* **2019**, *321*, 1820–1821. [[CrossRef](#)] [[PubMed](#)]
34. Nguyen, T.V.; Eisman, J.A. Post-GWAS polygenic risk score: Utility and challenges. *JMPR Plus* **2020**, *4*, 1–5. [[CrossRef](#)] [[PubMed](#)]
35. Lello, L.; Raben, T.G.; Yong, S.Y.; Tellier, L.C.A.M.; Hsu, S.D.H. Genomic prediction of 16 complex disease risks including heart attack, diabetes, breast and prostate cancer. *Sci. Rep.* **2019**, *9*, 15286. [[CrossRef](#)] [[PubMed](#)]
36. Patterson, N.; Price, A.L.; Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2006**, *2*, 2074–2093. [[CrossRef](#)]
37. Chang, C.C.; Chow, C.C.; Tellier, C.L.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **2015**, *4*, 7. [[CrossRef](#)]
38. Barrett, J.C.; Fry, B.; Maller, J.; Daly, M.J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **2005**, *21*, 263–265. [[CrossRef](#)]
39. Collister, J.A.; Liu, X.; Clifton, L. Calculating Polygenic Risk Scores (PRS) in UK Biobank: A Practical Guide for Epidemiologists. *Front. Genet.* **2022**, *13*, 818574. [[CrossRef](#)]